The Large Scale Structure of the Universe

Thesis by

ALAIN PICARD

In Partial Fulfillment of the Requirements

for the Degree of

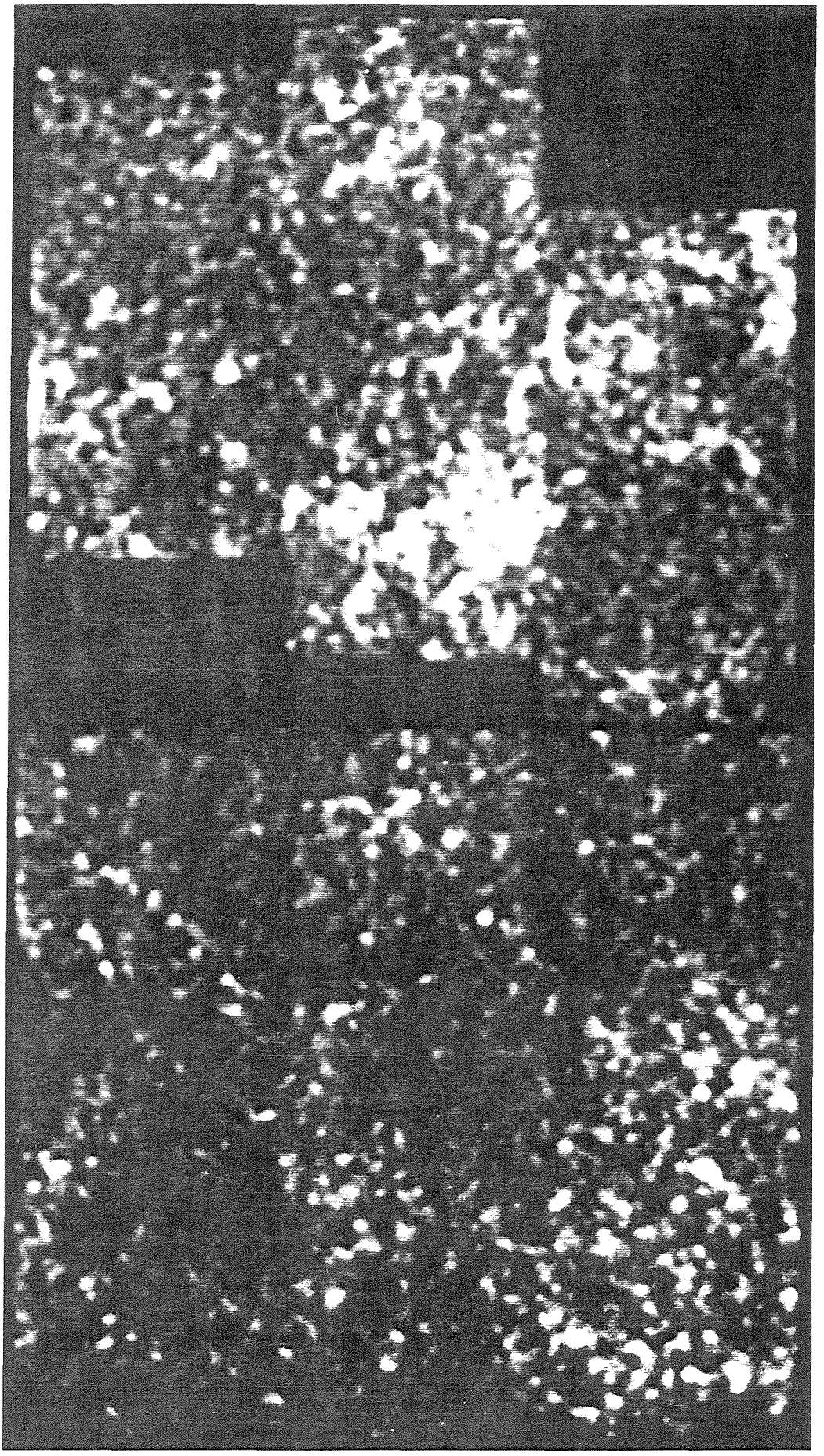
Doctor of Philosophy

California Institute of Technology

Pasadena, California

1991

(Defended May 22, 1991)



A ma mère,
qui, elle, a déjà tout compris...

Hâtez-vous lentement; et sans perdre courage,
Vingt fois sur le métier remettez votre ouvrage:
Polissez-le sans cesse et le repolissez;
Ajoutez quelquefois, et souvent effacez.

Boileau (1636–1711)
L'Art Poétique

Acknowledgments

When I first came to Caltech, I thought I was going to be a theorist. No, really! Just ask Chris Wilson, she'll tell you.

Well, as fate would have it, I eventually saw the light (pun intended) and became an observer. I'm not sure how it happened, but it must have had something to do with those many winter nights I spent at Palomar with Bev Oke, who perpetually needed someone to "run the keyboard" for him, while he was riding prime focus... I learned an enormous amount about observing from Bev, and for this and many other things I thank him very much.

I wish, of course, to give special thanks to my advisor, Jeremy Mould. Not only did he agree to take me on as a (somewhat disoriented) graduate student, but he then proceeded to let me transform the project he had suggested to me as a summer project into a full Ph.D. thesis, and gave me the leeway to do things pretty much the way I wanted to. The freedom with which he let me develop this thesis made it an especially valuable experience. How he always seemed to ask the right questions at each step still amazes me (I've *got* to learn how to do that!)...

This is the first thesis to come out of the second sky survey. The second POSS would not be possible without the generous contributions of the Samuel Oschin Foundation. I would also like to thank the Eastman Kodak Corporation, for providing Palomar Observatory with their excellent plates.

Money and equipment would be useless without the men and women who do all the hard work. As such, I would like to thank Bob Brucato and Bob Thicksten for their tireless efforts in making the sky survey a reality (and in Dr. Brucato's case, also for his work in keeping the Observatory as free of light pollution as possible). Furthermore, I would like to thank the "POSS" post-docs. namely Jim Schombert and Neill Reid, for the wonderful job they did (and which Neill is still doing) in making the sky survey a reality. I would like to include among these "heroes" Jean Mueller, for her help in acquiring my fields with high priority, and the great

enthusiasm which she displays in the survey, as witnessed by her record number of supernovae!

Having spent an enormous amount of time observing at Palomar, it is a pleasure to acknowledge the help of the men and women who make Palomar Observatory such a great place to work. The expert help of the night assistants, Juan Carrasco, Skip Staples and Will McKinley made observing a lot easier. And I cannot forget the incredibly competent engineers, Dave Tennant, John Hennings and Mike Doyle, who always managed to fix the problems just in time for observing.

In particular, though, I wish to thank Juan for our many pleasant conversations. When I think of Juan, there is one word that comes to mind: class. I admire his work and dedication enormously.

I wish to thank the Canadian government for financial help in the form of an NSERC scholarship, as well as Palomar Observatory for further financial support.

I wish to express my appreciation for Judy Cohen's work in improving the computing situation for all of us. When I first arrived, there were two VAX 11-780's, and that was *it*. Obviously, my thesis would never have seen the light of day under those conditions.

My thanks go to Michael Strauss, for many enlightening conversations, not to mention his excruciatingly detailed readings of my manuscripts, and his lending me his apartment for my post-thesis party. I thank Don Hamilton, for helping me set up FOCAS, and teaching me to appreciate Mahler and (ugh!) Bruckner. I will not rest until he discovers the joys of Josquin, Monteverdi and Ravel...

I wish to thank James Gunn and Jeff Hester with their help with the 60" PFUEI multi-slits, even though that work never made it into this thesis. Now *that* was hard work. As Jeremy says, after that, photometry is just a "leisurely activity for gentlemen..."

Everyone told me how life at Caltech was so awful, dull and boring. I'm afraid I must have missed something, for I found it nothing of the sort. Surely, this must be due to the camaraderie which developed between myself and my classmates, in

particular those of my year, Chris Wilson, Rich Rand, Mingsheng Han and Kent Budge. Maybe it's just that we had a different idea of fun, but I certainly remember fondly all our outings at the LA Phil... not to mention our PJ trips, and outings to Magic Mountain. But best of all, though, was the Friday tradition of Beer at the Ath, followed by endless bickering about whether we should go for Chinese or Mexican food.

I wish to thank John Apostolakis, Paolo Coppi, Julie Moses, and all those others who made softball such an *interesting* experience...

I wish to thank the students in the years before mine. I understand that Steve Myers, Chuck Steidel and Blaise Canzian had somewhat of a traumatizing experience; I think that they laid the ground work for making Caltech a more pleasant place to study. I like to think that watching Steve's thesis agonies helped me avoid them to some extent.

The most senior student with whom I really had a chance to interact is Jim McCarthy; I must say from the outset that I think of him as a hero riding off into the sunset. He and his programs taught me a great deal about FIGARO and good programming in general.

And life could not possibly be boring in the presence of the third year students, especially José Navarro, Nick Weir and Tom Herbig. I think the world of astronomy had better watch out! It is reassuring to see such deeply dedicated students still be so friendly and cultured; it convinces me that one need not be a nerd to be successful...

I would like to thank those students who shared my advisor: Chris Tinney, Josh Roth and Mingsheng Han. What a motley crew! Our diversities must be a testament to the great breath of Jeremy's scientific interests...

I wish to express my thanks to our librarian, Helen Knudsen, for her kindness and deep dedication to the library. The number of hours she puts in above and beyond the normal call of duty helps make the library a more pleasant and efficient one to use. My thanks also go to Diane Fujitani and Gita Patel. My talents for

dealing with paperwork are rather limited, as they can attest, and so I am doubly indebted to them for helping me remain oblivious to the "real world".

My thanks also go to Ann Palfreyman and Bob Brucato, who keep the Palomar Office running, in spite of us astronomers' profound lack of understanding in matters logistical.

I wish to thank Chris Kochanek and Barbara Wyszlouzil (Boris and Natasha) for many fine dinners, followed by recorder playing and excessive gluttony in the forms of outlandish cakes. But seriously, I am thankful for their effort in making it to my wedding, and being such excellent friends and companions.

I would like to offer my most heartfelt thanks for those people who gave me their support when Helen was in the hospital, and it wasn't clear that she was going to make it. I will never forget the hours spent in that small waiting room, where the atmosphere kept oscillating between the hilarious and the tragic. That's what great stress does to people. I have seen many families where there is not as much caring as I experienced then, and it made me realize just what a special group of people I was privileged to work with. I will not name names, but you all know who you are; thank you, for both of us.

I wish to thank my family, Maman, Marie-France and Georges, who have always supported me, and whose inordinate pride in me I take as the greatest compliment and gift I could ever be given.

Finally, there are two very special persons I wish to thank.

I wish to thank Jean Quashnock, my lifelong friend. It is fair to say that, if I had not met him, I probably would not have become a scientist. *Merci pour tout, Jean; on se comprends.*

And lastly, my thanks go to my wife, Helen Johnston. Her love, care, understanding, kindness and support have kept me going through the difficult times. Perhaps it is nearly losing her which has made me truly realize how precious she is to me. With her at my side, I can embrace life with open arms, every day. May I give her half as much as she gives me.

Abstract

We present a catalogue of galaxies from COSMOS scans of 16 second epoch Palomar Observatory Sky Survey (POSS-II) plates. The surveys covers an area of 386 square degrees, and covers 150,000 galaxies in the magnitude range $16.5 \leq r \leq 19.0$. The plates are divided into a north galactic (7 plates, $\alpha \approx 15^h\,0^m, \delta \approx +30°\,(1950.0), b \approx 65°$) and south galactic (9 plates, $\alpha \approx 0^h\,20^m, \delta \approx +5°\,(1950.0), b \approx -55°$) field.

The magnitude scale of each plate was brought to a common photometric system (Gunn r) using extensive CCD photometry. The total error in the zero point of the magnitude scale at the faint end is 0.06^{m}.

The procedure for separating stars and galaxies is based on image parameters provided by the COSMOS machine and is completely objective. Visual tests indicate that the catalogue is 94% complete.

We have measured the angular correlation function of galaxies $\mathrm{w}(\theta)$. The north field is in good agreement with the results of the APM survey. The south field exhibits more power on large scales than the north field. We have performed several tests to determine the effect of possible systematic errors between the two fields and do not find any effect sufficiently large to account for the south field's excess power. The measured correlation function for both fields are incompatible with the predictions of the currently most popular model, cold dark matter (CDM).

The number counts of galaxies in the two fields are substantially different. Several tests have been performed to make sure that the effect is not the result of a systematic error in zero point of the magnitude system or due to extinction or obscuration. The natural scale for structures which might be responsible for this difference is $50h^{-1}\,\mathrm{Mpc}$. We do not know if this difference is a rare event; more fields need to be surveyed to address this important question.

We have developed an objective algorithm for detecting clusters of galaxies from our galaxy catalogue. The algorithm is based on fixed overdensities at varying angular scales, and is well-suited to detecting clusters over a large range of redshifts.

We present a catalogue of clusters based on this algorithm, using an overdensity threshold obtained from comparing the cluster's densities to each field's overall background density. The results are very sensitive to the choice of background density. The multiplicity function of the objective clusters agree well when referenced to each field's background density, but disagree sharply when referenced to the overall mean density of the survey.

The objective catalogue is compared with that of Abell's. Poor agreement is found between the two catalogues. It is not possible to choose a set of parameters in our catalogue which reproduces Abell's catalogue well.

The measurement of the galaxy-cluster correlation function $w_{g-cl}(\theta)$ is in good qualitative agreement with predictions of CDM, but more detailed comparisons with theory must wait for the acquisition of redshifts for a large fraction of the clusters.

At present, no theory is able to explain the amount of power on large scale observed in this survey.

Table of Contents

1. Introduction

Large, uniform surveys, whether they be of stars, galaxies, quasars, or VLBI radio sources, have always held a very important place in astronomy. Without systematic surveys, which can be culled for meaningful trends, astronomy turns into botany, and any quantitative analysis becomes meaningless.

And of all surveys, it is fair to say that the Palomar Observatory Sky Survey (POSS) has been one of the most influential and scientifically useful ever. So much so, in fact, that it became evident that now, 30 years later, with the improvements in photographic emulsions, it was time for a new sky survey. This thesis contains some of the first † scientific results from the second epoch Palomar Observatory Sky Survey (POSS-II Reid *et al.* 1991).

Of equal importance to the improvements in photography, there have been tremendous improvements in the techniques for scanning plates, and on- and off-line computer processing to enable us to rapidly generate reliable catalogues of astronomical objects with accurate photometry and objective classification criteria. The two most prominent such operations at the present are the COSMOS facility, at the Royal Observatory, Edinburgh, and the APM (Automated Plate Machine) at Cambridge. These two machines are essentially equivalent, and both represent the current state of the art in fast scanning and analysis.

This thesis is about observational cosmology, a subject eminently suited to study through the analysis of large surveys. The fundamental problem we wish to investigate is a simple one: the distribution of matter on cosmologically interesting scales. More precisely, the physically interesting parameter is the spectrum of mass density fluctuations in the present universe. We now know that this is not directly observable because a large fraction of the matter in the universe is not luminous, and so what we must observe are *tracers* of the underlying mass distribution.

† Some work has been done by Schombert (1988) on low surface brightness galaxies, and by Tinney (1991) on low luminosity stars.

The most popular tracers are galaxies. The reasons for this are immediately apparent: they are fairly bright, and thus can be seen out to cosmologically interesting distances, and they are easy to recognize. Their observational properties are well-known, and they are reasonably well-understood physical systems. A successful theory, in fact, should be able to explain how galaxies are formed from the underlying mass density, and so it has become customary to use these objects as tracers, introducing the concept of the *bias parameter*, which simply measures how much more strongly (or weakly) they correlate than the mass distribution.

The core of this thesis is therefore the construction of a large, uniform catalogue of galaxies. This is the subject of Chapter 2. The catalogue presented is based on 16 plates from POSS-II. This makes the catalogue large enough (it contains 150,000 galaxies) to provide an adequate basis for statistical description. We believe that, at present, this is the best calibrated catalogue of galaxies in existence (although it is by no means the largest). We have attempted to give a sufficiently detailed description of the calibration procedure to enable another astronomer to duplicate the work exactly. The catalogue is divided into two regions of the sky, one in each galactic hemisphere, thus enabling us to see if our two regions of space behave similarly. They do not.

At present, theories of the formation of structures in the universe are being squeezed between two major pieces of observational evidence. At high redshifts, the smoothness of the cosmic microwave background (*e.g.* Meinhold and Lubin 1991) is in conflict with our current best guess of the primordial power spectrum. At low redshifts, the extreme "lumpiness" of the distribution of visible matter (chiefly, galaxies and clusters of galaxies) calls for a lot of power on large scales in the power spectrum (this work, and Maddox *et al.* 1989). Current theories are floundering on how to simultaneously satisfy both these constraints, *i.e.* on how to produce so much structure from an initially very smooth spectrum. The best theory at present, Cold Dark Matter (CDM, Bardeen *et al.* 1986) is in serious difficulty, as it is unable to explain the amplitude of $w(\theta)$, the two-point angular correlation function of galaxies,

on scales of $10h^{-1}$ Mpc[†], and if the correlation length of clusters of galaxies is indeed $25h^{-1}$ Mpc, as indicated by the work of Bahcall and Soneira (1983), this theory is certainly not correct. There is, at the time of writing, no better candidate theory.

The evidence for power on large scales is still controversial, as the measurements involved are extremely sensitive to even very small systematic errors. The measurement which, more than any other, has begun to signal the beginning of the end for CDM is the APM (Maddox *et al.* 1989) measurement of $w(\theta)$. Another group (Collins *et al.* 1989) is also measuring $w(\theta)$, and the result of these two groups is now converging[‡]. However, it must be stressed that *both* these groups are using *exactly the same photographic material*, and neither group has very much CCD data to calibrate the survey.

The degree of skepticism which these measurements of $w(\theta)$ have met is due to the crucial dependence on the survey's uniformity and the data reduction techniques employed. Geller *et al.* (1984) have shown that mismatches in a survey's uniformity of 0.05^{m} on 5° scales (the size of a survey plate) can cause a "break" in $w(\theta)$ identical to the one observed. In Chapter 3, we present our measurement of $w(\theta)$. The most important aspect of the work presented in this thesis is, to my mind, the fact that we confirm the presence of power on large scales based on a catalogue which is *completely independent* of previous work, both in plate material and calibration and reduction techniques. Furthermore, this power is seen in *both* of our survey fields. This is indeed good news for the observers, as Chapter 2 hints at some of the difficulties in calibrating large plate mosaics using the method employed by the APM and COSMOS groups.

In Chapter 4 we discuss number counts of galaxies. This is a much revered topic in astronomy, with a history going back (at least!) to Hubble himself. The counts,

† Throughout this thesis, we shall refer to cosmological distances in h^{-1} Mpc, where h is the value of the Hubble constant measured in units of 100 km/s/Megaparsec.

‡ Preliminary results of Collins *et al.* , based on a small region of the sky, were in disagreement with the APM's measurement. Recent results, based on a larger plate mosaic, now agree well with the APM (R. Nichols, *pers. comm.*).

being in a red passband, are not very sensitive to galaxy evolution, and are thus not of great interest in that sense. We might in fact have glossed over the galaxy counts if it had not been for a substantial discrepancy between the counts in the two fields. Large variations in galaxy counts in different regions of space are not new (see Maddox *et al.* 1990;) however, they had not been previously observed in surveys encompassing such a large volume of space. Again, it took a lot of tests to convince myself that the effect is indeed real, and not due to systematic effects. The clinching argument, in my opinion, is that it is not possible to *both* bring the galaxy counts into agreement *and* obtain star counts which agree reasonably well with Galactic models.

This chapter shows, then, that even when comparing two rather large volumes of the universe ($V \approx (125h^{-1}\,\mathrm{Mpc})^3$), we measure mean densities which are substantially different, and this is certainly evidence for "Large Scale Structure", although how large the scales are, and what the structure is, is at this point unknown. This is the topic of Chapter 4, and serves as an excuse for the grandiose title of this thesis.

Galaxies are abundant, which is both a virtue and a problem. Their great numbers mean that it is very time-consuming to systematically measure their properties in a cosmologically large volume of space. The largest redshift survey, (CfA survey, de Lapparent *et al.* 1988) has been ongoing for several years, and its volume is still not large enough to be representative of the universe as a whole. This is why it is also popular to use *clusters of galaxies* as a tracer. Being far less numerous, it is possible to obtain complete redshift information for samples of clusters in a reasonable amount of telescope time. Unfortunately, clusters suffer from severe problems of their own. We have even less idea what kind of bias parameter is needed for clusters than we do for galaxies. Worse: the very definition of a cluster is uncertain and problematical. Where most observers would agree on what is and what isn't a galaxy, they would certainly *disagree* on what exactly constitutes a cluster. If one imposes very stringent criteria, (say, by defining a cluster to be a physical

system containing so many galaxies, whose velocity dispersion is larger than a given number), one nullifies the positive aspect of using them as tracers, as the telescope time involved in obtaining the redshifts to ascertain the clusters' veracity becomes prohibitive.

In 1958, a graduate student defended a Ph.D. thesis at Caltech, presenting a thesis which has a lot in common with this one. I am referring of course to George O. Abell. I was not fortunate enough to meet† Dr. Abell before his untimely death in 1983, but because of the happy coincidence of our work (and choice of graduate institution) I have developed a profound respect for his work.

The Abell catalogue of galaxies is the one most frequently used today, both as a "list of nice clusters" (for which it is eminently well-suited) and as a basis for statistical analysis. This is a somewhat more touchy point. The lack of objectivity in Abell's selection procedure, and the possible non-uniformities of the catalogue due to plate-to-plate variations have been blamed for some of the results derived from the catalogue (in particular we are thinking of the large amplitude for the cluster-cluster correlation function). There is now hope of solving this problem with large, uniform surveys of galaxies such as the one presented in this thesis. From such surveys, it becomes possible to search for clusters of galaxies in a rigorous, objective manner, and photometric variations are taken care of by independently calibrating each plate. Many problems still remain (mainly selection effects and foreground/background contamination) but it should become possible to address these questions quantitatively. This could be done, for example, by comparing a two-dimensional cluster survey to an X-ray selected sample of clusters, and optimizing the search parameters to reproduce the X-ray survey, or through Monte-Carlo simulations to quantify the selection effects.

The main thrust of Chapter 5, then, is to

i) develop an algorithm to detect clusters from the galaxy catalogue,

ii) present the thus obtained catalogue, and

† Through a fortuitous set of circumstances, I did manage to inherit Abell's personal copy of the Zwicky *Catalogue of Galaxies and of Clusters of Galaxies*, which I cherish. Thanks Ann!

iii) compare this catalogue to that of Abell, in order to ascertain the degree of correspondence, which is relevant in light of the numerous statistical studies performed on Abell's catalogue. We then proceed to calculate the galaxy-cluster correlation function, which is shown to agree well with the predictions of CDM.

If some think that I am "dumping on Abell" in Chapter 5, I assure you that this is not so. I have had to examine closely only 18 plates; Dr. Abell worked with several hundred. The fact that only now, in the 1990s, are we able to do a better job than he did in 1958, should be a sufficient testimony to the endurance and quality of his work.

If this thesis is one tenth as successful as Abell's, I shall be delighted indeed.

1.1 References

Abell, G. O. 1958, *Ph.D. thesis, Calif. Inst. of Tech.*

Bahcall, N.A. and Soneira, R.M., 1983, *Astrophys. J.*, **270**, 20.

Bardeen, J. M., Bond, J. R., Kaiser, N. and Szalay, A. S., 1986, *Astrophys. J.*, **304**, 15.

Collins, C. A., Heydon-Dumbleton, N. H. and MacGillivray, H. T., 1989, *M.N.R.A.S.*, **236**, 7p.

Geller, M. J., de Lapparent, V. and Kurtz, M. J., 1984, *Astrophys. J. (Letters)*, **287**, L55.

de Lapparent, V., Geller, M. J., and Huchra, J. P., 1988, *Astrophys. J.*, **332**, 44.

Meinhold, P. and Lubin, P., 1991, *in press.*

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989, *M.N.R.A.S.*, **242**, 43p.

Maddox, S. J., Sutherland, W. J., Efstathiou, G., Loveday, J., and Peterson, B. A., 1990, *M.N.R.A.S.*, **247**, 1.

Reid, I. N., Brewer, C., Brucato, R. J., McKinley, W. R., Maury, A., Mendenhall, D., Mould, J. R., Mueller, J., Neugebauer, G., Phinney, J., Sargent, W. L. W., Schombert J. and Thicksten, R., 1991, *The Second Palomar Sky Survey*, submitted to *P.A.S.P.*

Schombert, J., 1988, *Astron. J.*, **95**, 1389.

Tinney, C. and Mould, J. R., 1991, submitted to *Astron. J.*

2. The Galaxy Catalogue

2.1 Introduction

In this Chapter, we describe the procedure used to establish our catalogue of galaxies. What requirements would an ideal catalogue satisfy? First, the catalogue should be *large*, since we wish to survey as large a volume of space as is practical. This is necessary because of the nature of the problems which we wish to address using the catalogue, namely the determination of the clustering properties of galaxies and, eventually, clusters of galaxies. Second, and equally important, the catalogue should be *uniform*, in its most stringent sense that the probability of detection of a galaxy should depends only on the intrinsic properties of that galaxy (distance, luminosity, morphology *etc.*) and on no parameters extrinsic to the galaxy, *e.g.* its position on the plate, which plate it happens to be observed on, *etc.*

To satisfy both these criteria is an extraordinary task, and it must be realized that the two criteria are at odds with one another. A very deep CCD image on a photometric night will easily satisfy the uniformity criterion, for example, but will only contain a few hundred galaxies, in the best case. A catalogue visually compiled from photographic material will not satisfy the second criterion very well, as visual classifications are difficult to keep uniform.

There are two ways to satisfy both constraints: a large photographic survey scanned on a photodensitometer, or a large CCD transit (or even "point and shoot") survey.

The second method is preferable because of the great stability of CCD detectors, but requires a dedicated telescope. Such surveys are in fact planned for the future. The first method, which requires extensive CCD calibration to compensate for the variable nature of photographic photometry, is the one followed in this thesis. Even here, it must be realized that the cost in telescope time is prohibitive, and I was very fortunate to have the opportunity to undertake this project at Caltech.

Our survey is still not very large, but large enough to probe an interesting volume of space, or roughly $(125h^{-1}\,\mathrm{Mpc})^3$. This is as large as we could make it

and still satisfy the real world constraint of available telescope time. The completed galaxy survey consists of some 150,000 galaxies, with accurate photometry on a common system, and with the effects of plate to plate variations removed as well as they can be. The total area surveyed is 386 square degrees.

The only surveys comparable to ours are those obtained by the APM (Maddox *et al.* 1989) and the COSMOS group (Collins *et al.* 1989), both groups using the United Kingdom Schmidt Telescope (UKST) plates of the southern sky. The APM survey, in particular, covers 4200 square degrees, more than 10 times the solid angle of our survey. The photometric calibrations of this survey, however, are based on internal consistency rather than being externally imposed by CCD sequences. As we will see in this chapter, we have evidence that this may be a dangerous technique.

The galaxy catalogue presented here reflects the current state of the art, and is the most extensively calibrated to date. As such it is suitable for a wide range of projects. We have carried out some of these in this thesis, and will discuss future work based on this data set in our conclusion.

2.2 Plate Material

The plate material for this project consists of 16 IIIaF plates obtained at the Oschin 48" Schmidt telescope in the course of the second Palomar Observatory all Sky Survey (POSS-II Reid *et al.* 1991). The IIIaF plates are exposed, typically for 60 minutes, through an RG610 filter, and are well suited to detecting the typical galaxies found in clusters (red ellipticals).

These plates cover two regions of the sky: our northern field ($\alpha = 15^h, \delta = 30°, b = +65°$) and our southern field ($\alpha = 0^h20^m, \delta = +5°, b = -55°$). The fields were to have been squares of 3×3 plates, thus subtending an angle of approximately 15° on a side. Unfortunately, two of the plates, both in the north field, were found to be of unacceptable quality (Fields F385 and F512) and had to be rejected, so the north field is in fact not square, and covers a smaller area than the south field.

These fields were chosen primarily because *i*) they are at high galactic latitude, where Galactic obscuration is not likely to be a large problem, and *ii*) the best months to observe them from Palomar are May and October, when the weather is reasonably favourable. We chose to survey two regions of the sky, rather than one larger region, so that we could compare our results in the two areas, and also so that CCD observations could be done twice a year, minimizing the likelihood of being completely unable to acquire data due to bad weather.

When more than one plate existed in a field of interest, the plates were visually inspected and the best one was chosen for scanning. Unfortunately, the plates vary in quality (mainly in limiting magnitudes), so we are limited by the worst plate if we are to keep the galaxy sample homogeneous. Details of the plate material are given in Table 1 of Chapter 4.

2.3 CCD Photometry

2.3.1 Observations

We first inspected each plate visually with a 8× lens and looked for clusters or groups of galaxies; we typically found about 30 such objects per plate. We gave each object a number, so for example, F386Cl34 is the 34^{th} object we circled on field 386. We performed rough astrometry (good to about 5″) and gave a rough quality grade to each cluster; from poor and doubtful to rich and spectacular. On the new survey plates, the Abell clusters usually stood out very easily, but it is especially the unknown, often faint and compact clusters that interested us. This preliminary catalogue of clusters of galaxies is thus severely biased and incomplete; its purpose was simply to find good regions to observe in order to perform the plate calibration.

We obtained photometric data for the Abell clusters and as many of the clusters identified by eye in the fields of interest as time permitted. These data consisted of pairs of images obtained in the Gunn g and r bands, with typical exposure times of 1800 s in g and 1200 s in r. These data were obtained at the Palomar 60" telescope using CCD cameras 4 and 8. The detectors in these cameras are charge coupled devices manufactured by Texas Instruments and have an array size of 800 × 800, with a pixel size of 15 μm. The chips have a gain of approximately 2 electrons per DN, and readout noise of 9.5 electrons. The exposure were long enough that we were always well above the readout noise, *i.e.* , we are sky-limited. The data were acquired with the re-imaging lens in place to obtain a suitable field of view; this system yields a pixel size of 0.615″ and a total field of view of 8′.

This pixel scale is just small enough to enable us to determine the point spread function for the stellar images reasonably well, a step necessary for the star-galaxy separation algorithm of FOCAS. We noticed, however, that the images suffer from coma at the extreme edges of the field of view (this is a defect of the re-imaging system), so we only used images detected on the central 760 × 760 pixels to ensure uniform classification and photometry. Data were acquired on every night on which weather permitted; however only the data obtained on nights which were judged to be of photometric quality were used for calibrating plates.

2.3.2 Reduction of Photometric Standards

Standard stars defining the Gunn $g\,r\,i\,z$ photometric system (see Kent 1985, Thuan and Gunn 1976) were observed. Typically, 5 or 6 stars were observed each night. Each observation actually consisted of a multiple exposure where we moved the telescope between each snapshot; this way we effectively obtained three observations for the price of one. We observed stars spanning a wide range of colors, to more accurately determine the color coefficients of the photometric equations. The data were then reduced as follows.

First, we flat fielded the data using the standard procedure of subtracting the bias level and dividing each image by a suitably normalized exposure of the top of the dome uniformly illuminated by an incandescent light source. Some flat fielding problems† were experienced, especially in the r band; a strange "doughnut" shape, presumably caused by the secondary mirror and its support structure, would appear in the flattened images. The problem was most severe at the beginnings of the nights, suggesting some sort of thermal equilibrium problem, and so flat-fields were taken after each r observation, in order to minimize the problem. To save time, the telescope was left in place and the dome swung so that the flat field lamp (which is mounted behind the secondary mirror) was pointing at the dome. Also, we felt that if this was a flexure problem in the camera assembly, it was best to not move the telescope between the program exposure and the flat field. Fortunately, this anomaly only affects a small fraction of the pixels on the frames, and is at low level (1 or 2%) and so was ignored in subsequent reductions, with the hope that it would essentially only increase the scatter in the magnitude calibration curves.

We then used the FIGARO (Shortridge 1984) astronomical reduction package to determine the instrumental magnitudes of the standard stars. We performed aperture photometry on the standard stars using the programs CENTERS and FOTO. The algorithm used by FOTO is quite straightforward: all the counts in the pixels lying within a user-specified radius from the centroid (determined by CENTERS)

† This strange problem had disappeared as of Fall 1989. To the author's knowledge, no one at Palomar Observatory knows why, or what caused it in the first place.

are added up, and a value for the sky is obtained by averaging the data values in an annulus of radii also determined by the user. The radial profile of the object is displayed on a graphics device to allow the user to choose these radii in an optimal fashion. When doing this, we were very careful to select apertures sufficiently large to enclose all the light of the star. These apertures were typically 6″, except in the cases where the standards were so bright (*e.g.* HD171164, HD84937) that they had to be observed out of focus. There the apertures were usually 13″.

2.3.3 Photometric Transformations

After having obtained the instrumental magnitudes g_i and r_i for each star, we used the method of least squares to find the best fitting line of the forms (1) and (2) to the data, as shown on Figures 1 and 2.

$$g - g_i + k_g \sec(z) = A_g + B_g(g - r) \tag{1}$$

$$r - r_i + k_r \sec(z) = A_r + B_r(g - r) \tag{2}$$

where the A's and B's are the coefficients to be determined and z is the airmass at which the observation was made. The values used for the extinction coefficients were taken to be $k_g = 0.18$ and $k_r = 0.10$ (Kent 1985).

Rather than determining the slopes B_g and B_r independently for each night, all the data which were judged photometric (as determined both by the values of the scatter in the plots, and the careful notes kept in the observer's log book) were collated together and a mean slope for the entire run was determined. This was done because we do not expect these coefficients to vary from night to night; however the zero point coefficients A_g and A_r were determined on each night, to allow for the possibility of differences in atmospheric transparency during each night of a given run. Typical values of the A's and B's are shown on Figures 1 and 2.

The formal uncertainties on these quantities are also shown, as is the value of the *rms* scatter about the best fit line. The data were considered to be of photometric quality if this *rms* value was not larger than 0.035^{m}. Several times during each night,

we assessed the quality of the weather, and made a note of it in the observation logbook. We were pleased to find a very strong correlation between these notes and the dispersions in the photometric reductions.

2.3.4 FOCAS Magnitudes

The CCD observations of clusters of galaxies were reduced using FOCAS (Faint Object Classification and Analysis System, Valdes 1982, Jarvis and Tyson 1979). FOCAS is a set of programs to generate and analyze catalogues of astronomical objects from raster images, such as PDS scans or CCD data. FOCAS is particularly suited to performing photometry on faint galaxies, unlike, for example, DAOPHOT (Stetson 1987) which is geared towards stellar work.

The main programs of FOCAS are:

- *sky* and *skyavg*, which determine the sky background to be subtracted from each pixel,
- *detect*, which finds all the images present on the frame,
- *findstars*, which finds objects suitable for use as templates to determine the point spread function on the frame,
- *splits*, which raises the detection threshold to determine if any given object is actually a merger,
- *evaluate*, which calculates the actual magnitudes and other image parameters (such as aperture and isophotal magnitudes, second moments, *etc.*), and
- *resolution*, which does the star-galaxy classification, discussed in § 2.6.1.

FOCAS provides both aperture and isophotal magnitudes. Since the COSMOS magnitudes are isophotal, it was decided to use isophotal magnitudes to calibrate them. The isophotal threshold used in the magnitude calculation is given to FOCAS as a multiple of the sky variance on any frame. After examination of several frames, we adopted a "standard" sky brightness for the g and r bands *i.e.* , a typical sky

brightness at Palomar measured on our CCD frames. These values are, respectively,† 20.80 r mag/["]^2 and 21.20 g mag/["]^2. We then calibrated one of our fields (Field 449) with three isophotes for our magnitudes (isophotes of 1.5%, 3% and 7% of our "standard" sky) to determine which isophote gave the tightest correlation of the FOCAS isophotal magnitudes with the COSMOS magnitudes. The 3% isophote was marginally better than the other two, so it was adopted. These isophotes therefore correspond to $r_{3\%\mathrm{sky}} = 24.61$ mag/["]^2 and $g_{3\%\mathrm{sky}} = 25.01$ mag/["]^2.

The *findstars* program is actually a UNIX shell script which attempts to find bright, isolated stars by taking objects near the mode of the distribution of second moments of the image distribution. This algorithm is not robust, and will often try to use images near a bad or saturated column. So we inspected all of the point spread functions, and when they did not look right, (being too elongated, asymmetric or not monotonic, for example) we examined which stars it had used as templates. We then removed the offending objects, and re-computed the point spread function.

2.3.5 Limiting Magnitude of CCD the Photometry

Although the limiting magnitudes of the CCD frames vary from frame to frame, due to differences in transmission, exposure times *etc.*, all our CCD observations reach a limiting magnitude which is much fainter than that of the photographic plates. Figure 3 shows histograms of number of galaxies observed as a function of magnitude for four clusters (in both colors). Based on these data, we judge the completeness limit on the CCD frames to be 22.0 in r and 23.0 in g. Of course, these numbers should be regarded as guidelines, since as already mentioned, these numbers vary from frame to frame.

We will see later that the magnitude limit of the plates range from 20.5^{m} to 21.6^{m}. The exposure times were chosen to yield a reasonable signal to noise ratio for the faintest objects of interest, *i.e.* , those at the limiting magnitude of the deepest

† We denote the unit of surface brightness, magnitudes per square arc-seconds, by the symbol mag/["]^2.

plate. Going much deeper is not desirable, as the brighter galaxies become saturated (and hence useless for calibration) and not enough frames can be obtained.

2.4 COSMOS Data

2.4.1 General Description

The plate material was scanned using the COSMOS (Co-ordinates, Sizes, Magnitudes, Orientations and Shapes) machine at the Royal Observatory, Edinburgh (ROE). This scanner is described extensively in Macgillivray and Stobie 1984, so only a brief summary will be given here.

The machine itself is a high-speed scanner of flying-spot design. A high intensity beam of light from a cathode ray tube is split into two; part of the beam goes to a reference detector, the other part through the plate to a second detector. The plate density-to-intensity transformation for each pixel is performed by a lookup table generated from a spline fit to the density-to-intensity table obtained from the sixteen calibration spots at the bottom right of each survey plate. The data for the lookup table were not available for our plates, however, so a table was picked arbitrarily for a UKST plate which had the same background density as our plates. This error causes non-linearities in the magnitude scale, but these are removed by our procedure for calibrating the plates using CCD photometry.

A quick pass is performed over the plate to determine the overall background intensity (sky values). This is done by subdividing the plate in 64 by 64 squares, and in each of these calculating the mode of the transmission values. This matrix of numbers is used to predict smoothed transmission values for the background which are then used to calculate the amount of sky to be subtracted at each pixel. A contour plot for this matrix is shown on Figure 4. Note the gradients near bright stars, and the overall falloff near the edges and corners of the plates, caused by vignetting. The referencing of the magnitudes to 7% of the *local* sky is an attempt to make these magnitudes independent of their position on the plate; however these "field effects" can still cause systematic errors in the magnitudes. For the UKST plates, for example, the *area*–COSMAG locus moves when comparing images located

in the center of the plates to those near the edges. Such a gross effect has not been detected on our plates (although more subtle effects will be discussed later). This difference may be attributable to the fact that the POSS-II plates undergo nitrogen gas flushing during the exposure, which helps eliminate progressive desensitization of the emulsion by air leaking into the plate holder.

The magnitudes are determined by summing the (sky-subtracted) intensities for all the pixels in an image that are above the threshold cut, and thus are isophotal magnitudes. For our data we used a cut of 7% of the sky, which easily finds all images of interest, yet is high enough that very few noise or fake images are introduced. These are mostly produced near scratches in the emulsion or near the halo of very bright stars. Figure 5 shows the COSMOS output near a typical bright star, showing the diffraction spikes and numerous junk images in the halo. These are removed by hand by simply flagging and deleting the offending regions of the catalogue.

The final output from COSMOS consists of a set of 32 parameters for each image detected. The most important parameters are listed below:

Table 1. COSMOS Parameters of Interest

#	Name	Meaning	Units
1	*ra*	Right Ascension (1950.0)	decimal hours $\times 10^7$
2	*dec*	Declination (1950.0)	decimal degrees $\times 10^6$
7	*area*	area of the image	pixels
8	*ipeak*	maximum intensity of the image	intensity levels
9	COSMAG	COSMOS magnitude	mag $\times$ 100
10	*isky*	sky intensity	intensity levels
11	*xcen_i*	intensity weighted x centroid	0.1μm
12	*ycen_i*	intensity weighted y centroid	0.1μm
16	*a_i*	intensity weighted semi-major axis	0.1μm
17	*b_i*	intensity weighted semi-minor axis	0.1μm
18	*theta_i*	intensity weighted orientation	degrees
19	$\phi^\dagger$	classification parameter	
20	*pa*	position angle	degrees
29	*blend*	deblending flag	
31	*r*	calibrated magnitude	mag $\times$ 100
32	$\mathcal{S}^\dagger$	$\mathcal{S}$ parameter	

For our purposes, the most important of these are the positions, areas, magnitudes and peak intensities. The absolute positions $(\alpha_{1950}, \delta_{1950})$ are accurate to about 1.5″. The relative positions are much better than this; they are easily good enough, for example, to set up Palomar Observatory's multi-fiber Norris spectrograph, which requires an astrometric accuracy of 0.5″ over a 20′ field of view.

2.4.2 Deblending Software

The scans of the plates were processed using the COSMOS group's new "deblending" software (Beard and Thanish 1990), which processes the data twice: after

the initial determination of image parameters, each image (called the *parent* image) is re-thresholded at 8 higher surface brightness levels, and the resulting profile is examined to determine if the image is actually a blend of several images (called *daughter* images). This step is of the utmost importance for this project, as the galaxies in a cluster are often sufficiently close together that their isophotes merge at the isophotal cut chosen for detection, so the entire cluster can look like one large image. Clearly, if we are going to look for overdensities of objects, the clusters need to consist of several images! Figure 6 shows this process; the very large ellipse is COSMOS' best fitting ellipse to the "image" above the detection threshold (*i.e.* , most of the cluster). The smaller ellipses are those fit to the daughter images. When we compare this to a CCD image of the same cluster we find that about 80% of the images are recovered. It is mostly the very faint images which are not properly deblended. Fortunately, these would not have been included in the final catalogue anyway. We also note that the parameters for the deblended images are in general not very accurate, which causes these images to be classified as galaxies (the procedure used for classification will be discussed shortly). This is not a serious problem, as we prefer to accept a certain amount of contamination by stars (false detections) to throwing away real galaxies (incompleteness), especially those near the cores of clusters.

2.4.3 Accuracy of the COSMOS Magnitudes

The field centers of the POSS-II plates are separated by 5°. With the field size being 6.5°, this gives a 1.5° strip of overlap on each side. However, COSMOS only scans the central 5.3° of a survey plate, so our overlaps are in effect only 20′ wide. This overlap allows us to check the accuracy of the COSMOS magnitudes as follows. We can plot the values of any given parameter for each object present on both plates against each other. The scatter in these plots is a fair measure of the uncertainties due to both the scanning procedure and variation in the quality of different plates.

Figure 7 shows a dot plot of uncalibrated (raw) COSMOS magnitudes of all objects found on both plates in the overlap region between fields 448 and 449. Note

that, at bright magnitudes, there are two distinct loci of points. This is because the magnitudes for stars are more subject to effects of saturation of the emulsion, which are not repeatable from plate to plate, combined with the limited dynamic range of the COSMOS machine. This is why we have only used objects classified as galaxies to do the actual magnitude calibration. The scatter about the line which best fits the points below 1900 in Figure 7 is 0.15^{m}, indicating that the magnitudes are repeatable to about 0.1^{m} (since there is scatter in each direction).

2.5 Magnitude Calibration of the Plates

The magnitude calibration of the plates was done as follows.

For each CCD frame, we extracted from the relevant COSMOS catalogue all images lying within 10′ of the position of the center of the CCD field (as given by the header in the image file) to a small COSMOS catalogue. The position of the CCD frames on the plates is indicated on Figures 8 and 9, where the solid lines denote the boundaries of the survey and the dashed lines show where the individual plates were abutted. Each square represents a CCD sequence taken under photometric conditions. We then determined the transformation between the COSMOS *xcen_i* and *ycen_i* units to the (X,Y) pixel coordinates of the CCD data. The relevant data for these objects, namely their right ascensions, declinations, X and Y coordinates in the newly determined CCD system, as well as COSMOS magnitudes and ϕ parameters were then written to a file in the regular FOCAS catalogue format. (The FOCAS i/o routines are straightforward and well documented, so it was easy to write a program to do this[†]). Next, we used the FOCAS program *match* to generate a matched catalogue from the *g*, *r* and COSMOS images. This enables us to generate lists of magnitudes for corresponding objects and to easily apply any desired transformation to them.

After this had been done for all the photometric CCD data, we collated together all the data pertinent to a given plate, and plotted the calibrated FOCAS magnitudes of the objects against COSMOS magnitudes for all objects classified as galaxies by FOCAS, (stars were not used in this procedure because of the saturation problem mentioned in § 2.4.3 and illustrated in Figure 7). Figure 10 shows the calibration curve used for Field 752. Note the strong non-linearity of this relation, already discussed in § 2.4.1, due to the intrinsic properties of the photographic plates and our improper density-to-intensity lookup table.

The solid line shows a least square fit of the third order. The residual scatter is approximately 0.18^{m} for bright objects $r < 19.0$ and gradually gets worse at fainter

† Easy, that is, after this FORTRAN aficionado spent a few days learning the rudiments of the C programming language!

magnitudes. The different symbols represent objects selected from different CCD frames. Note that there is no obvious segregation of symbols (*e.g.* , there is no tendency for the squares to be systematically higher than the triangles). This is an important test of two things: first, that our photometric calibrations are well done (the CCD observations come from four different nights) and second, that the plate is *uniform*, that is, the magnitudes obtained from COSMOS are not dependent on the position on the plate.

2.5.1 Accuracy of the Zero Points

The error in measurement for individual galaxies near $r = 19$ is $0.15^{\rm m}$. However, it is the *systematic* errors which lead to spurious density fluctuations, and those are the ones we seek to minimize. As noted in § 2.3.3, the error in the determination of the zero point for the CCD photometry is $0.035^{\rm m}$ or smaller. Each plate was calibrated using four or more CCD fields. Furthermore, these were taken on different nights, so as to minimize possibility that a bad night (which for some reason still had a small dispersion in the photometric relationship) could cause problems. Since these are four statistically independent measurements, the error in the zero point of the magnitude scale of the plates *due solely to uncertainties in the calibration of the CCD material* is $0.035^{\rm m}/\sqrt{4} = 0.0175^{\rm m}$.

This error is minuscule, and so, in fact, the most important source of uncertainty in the zero point of the magnitude scales is that due to the fitting of the calibrated FOCAS magnitudes to the COSMOS magnitudes. Because the galaxy catalogue will be truncated, the relevant quantity is the uncertainty of the fit at the truncation limit. This uncertainty was estimated by fixing the linear and quadratic terms to their best values, and fitting the data with only the zero point term left free. The covariance matrix then indicates that the uncertainty in the zero point is $0.05^{\rm m}$. Thus the total error in the zero point of the plates is $0.06^{\rm m}$. We emphasize that only objects classified as galaxies by FOCAS were used in this calibration. Our formal error in the fit would certainly diminish if we include stars, but because of possible systematic differences in the COSMOS magnitudes between stars and galaxies mentioned in § 2.4.3, we have chosen not to include the stars.

Figure 11 shows a dot plot of magnitudes independently calibrated using the CCD photometry for objects in the overlap regions classified as galaxies between fields 448 and 449. We actually plot M(449)–M(448) so that we can expand the scale. The points should lie along the $y = 0$ line if the calibration is perfect. Most of our overlaps do not in fact obey this relation as well as the one shown in this figure; this problem will be discussed fully in § 2.9.

2.5.2 Magnitude Limits of the Plates

We have determined the magnitude limit for each plate by building a histogram of the number of objects (stars and galaxies) for each plate as a function of the calibrated magnitudes, as shown on Figure 12. The last bin before the number of objects drops is defined to be the magnitude limit of the plate. These values should be accurate to 0.1^{m}, and are only used to give us an indication of the quality of each plate. The magnitude limits are tabulated in Table 1 of Chapter 4.

2.6 Separation of Stars and Galaxies

2.6.1 Star Galaxy Separation using FOCAS

The star-galaxy separation on the CCD data was performed using the *resolution* classifier incorporated in FOCAS. This algorithm is fully described in Valdes 1982, and a complete description is beyond the scope of this thesis. Suffice it to say that the program employs a maximum likelihood method to fit the light distribution of each image to three templates which represent stellar, resolved and noise images, and then applies Bayesian statistics to decide which template is the best fit.

There are cases where the classification for an object on the g and r frames differ. When this happens, the object is always stellar. This is easy to understand, as small errors, cosmetic defects, cosmic rays, *etc.* can cause a star to be classified as a galaxy, but not vice versa. We thus classified objects as galaxies if they were classified as such on *both* frames by the automatic classifier, and as stars otherwise. Obviously the reliability of this automatic classification is greatly improved by having two CCD frames to work with, instead of just one. These classifications were tested against my own visual classification for several frames. The automatic classifier turns out to be excellent, in that I agreed with its classifications in more than 90% of the cases. The "misclassifications" fell into two major groups: objects that were so faint that I couldn't really decide whether the object was a star or a galaxy, and bright saturated stars which "bled" on the CCD image and so had definite non-stellar appearances. Other misclassifications occurred where an object fell on a cosmetic defect on the detector, such as the edge of the coronene blob on CCD 8; these errors alway go in the sense of classifying stars as galaxies, never vice versa. The galaxy catalogues generated by FOCAS are therefore essentially *complete*, but are contaminated by some stars.

Three fields were studied with particular care to quantify these effects: CCD frames F386cl34, F449cl09 and F513cl31. On these fields, 5 galaxies were missed by FOCAS out of 205 (or 97.5% completeness) and 10 stars were classified as galaxies out of 215 objects classified as galaxies (or 4.6% contamination). It is important to

remember that these statistics are for objects present on both the g and r frames *and* the COSMOS frame. The misclassifications increase for fainter and fainter objects, until nothing can be said for objects at the magnitude limit of any given observation. This is why we made sure that our CCD frames went at least 1^{m} deeper than the plates; we wanted to ensure reliable CCD calibration for objects as faint as the limiting magnitude of the photographic plates.

2.6.2 Photographic Star-Galaxy Separation

The star-galaxy separation on the plates is based on the fact that stars, being determined uniquely by the instrumental point spread function, occupy a well determined region in the measured parameter spaces. Galaxies, however, being resolved objects, have their own intrinsic profiles and thus tend to deviate from the region occupied by stars. As a simple example of this, consider a plot of the measured area of the objects against their instrumental magnitudes, as shown on Figure 13. At a given magnitude, (*i.e.* , at fixed $\log \sum_{pixels} I$) resolved objects, such as galaxies, have lower surface brightness than stars, hence a larger area, which causes them to lie in the upper regions of the diagram. This behaviour is typical, in that plotting any two parameters intrinsic to the image structure tends to show a well concentrated *stellar ridge*, and objects deviating from it which are galaxies, blends and emulsion defects.

The idea is therefore to separate stars from galaxies by "cutting up" portions in these various parameter spaces. Note however how quickly all objects merge in one ill-defined ridge for faint objects. This behaviour will occur for any set of parameters we choose, as it is simply a reflection of the fact that at the plate limit, stars and galaxies become indistinguishable, by eye or otherwise. We therefore have one last condition: the method for cutting up the parameter space should be objective.

The star-galaxy separation algorithm then goes as follows. We define, as done in Collins and Heydon-Dumbleton (1989), a quantity which we call the $\mathcal{S}$ parameter to be

$$\mathcal{S} = \frac{area}{\log(ipeak + isky) - \log(isky)(1 + \frac{cut}{100})} \tag{3}$$

where *area* is the area of the image (in pixels), *ipeak* is the peak intensity recorded for the object, *isky* is the intensity of the sky, and *cut* is the threshold cut above the sky (7% in our case). Physically, $\mathcal{S}$ measures the "hardness" of an image; the more centrally concentrated (small *area* and large *ipeak*), the smaller the value of $\mathcal{S}$. A plot of $\mathcal{S}$ *vs.* COSMAG for Field 449 is shown on Figure 14. Note that the separation here is cleaner to a fainter magnitude than when simply using the areas (Figure 13). The improvement, obviously, comes from our use of the extra parameter *ipeak* (*isky* is not really an independent parameter intrinsic to the object, and its introduction in Equation 3 is because we wish to reference the peak intensity to the local sky, which varies quite a bit over a typical plate.) For very bright objects, the limited dynamic range of the emulsion and scanner causes the denominator in Equation 3 to saturate, while the area of the image keeps increasing, accounting for the sharp increase of the stellar ridge for objects with COSMAG < 1500 on Figure 14. Also note the "stripiness" for the very faint objects, which is due to quantization noise. The last stripe is due to objects with an area of 4 pixels, the next to last 5 pixels *etc.*

Our next step is to obtain an objective characterization of the stellar ridge as follows. First we bin the $\mathcal{S}$ values in small magnitude bins in COSMAG ($0.2^{\rm m}$); this yields a histogram of the number of points having a given $\mathcal{S}$ value as a function of $\mathcal{S}$ for fixed magnitude, as shown on Figure 15. This figure is quite instructive: we see the sharp peak due to the stellar ridge, but we also see the large tail of objects (the galaxies) extending to large values of $\mathcal{S}$. It is clear from the broad shape of this tail that it extends well into the stellar ridge, and thus there is no hope of being complete.

We estimate the mode of this distribution using an algorithm described in Press *et al.* (1986). This algorithm goes as follows: we sort the $\mathcal{S}$ values in increasing order, yielding $(\mathcal{S}_1, \mathcal{S}_2, \ldots \mathcal{S}_N)$. We then choose a "window", say N/10 and compute $p_i = \mathcal{S}_{i+\rm window} - \mathcal{S}_i$ for all i. The value $\mathrm{M}_i = \frac{1}{2}(\mathcal{S}_{i+\rm window} + \mathcal{S}_i)$ for which p_i is a minimum is a good estimate of the mode.

We also estimate the width of the distribution by setting $\mathrm{W}_i = \mathcal{S}_i - \mathcal{S}_{5\%}$. ($\mathcal{S}_{5\%}$

is the $\mathcal{S}$ value of the 5^{th} percentile of the sorted values.) We use this value, as opposed to $\mathcal{S}_1$, to obtain a robust estimate of the mode. This works well, as there are always several thousand objects per magnitude bin. This yields a set of modes and widths (M,W) for each magnitude interval. These are plotted in Figure 16. Note how, when overlaid with the data points on Figure 14 the modes and widths delineate the star ridge to high accuracy.

We now fit a smoothing spline† (*not* a regular spline, which would force the curve to go through the knots defined by the data points) to the modes and widths. The spline fit to the modes is shown on Figure 17. The same is done to the widths. This is more difficult, as the fitting routine needs an estimate of the error bars for each point. For the modes, we simply give the routine the widths. For the widths, however, we have to make something up, so we give them $\sqrt{N}$ noise. Because of the quantization noise on Figure 14, the estimate for the width is sometimes inaccurate for the faintest magnitudes. Physically, the width of the stellar ridge should be a monotonically increasing function of magnitude, and so, when this is not obeyed, the offending values are removed by hand. We denote these two splines as M(COSMAG) and W(COSMAG).

Finally, we are in a position to define our star-galaxy separation criterion ϕ:

$$\phi = 1000 \times \frac{(\mathcal{S} - \text{M(COSMAG)})}{\text{W(COSMAG)}} \tag{4}$$

Note carefully the definition of ϕ: it measures *how far away* from the stellar ridge an object is located, *in multiples of the width of the stellar ridge at the appropriate magnitude.* The factor of 1000 is only put in because all these numbers are stored as integers, and we do not want to lose resolution.

A plot of ϕ *vs.* calibrated magnitude is shown on Figure 18. We have now encapsulated all that we know about the location of the object in the parameter space used to separate stars and galaxies in *one* number. Note that the stellar ridge is defined by $\phi = 0$, and that its width is essentially independent of magnitude.

† We used the routine ICSSCU from the IMSL set of mathematical routines.

2.6.3 Choosing the $\phi_{\rm cut}$'s

We must now, for each plate, choose a value of ϕ above which to accept objects as galaxies, and below which to reject them as being stars. We denote this number $\phi_{\rm cut}$.

We will set $\phi_{\rm cut}$ to be three times the width of the stellar ridge. Now recall that, by definition, the stellar ridge is defined by the equation $\phi = 0$. Because galaxies (with positive values of ϕ) tend to crowd the stellar ridge, especially at faint magnitudes, we estimate the width of the stellar ridge by computing the quantity

$$\sigma = \sqrt{\sum_{\phi \leq 0} \phi^2 / N(\phi \leq 0)} \qquad (5)$$

It will be immediately recognized that σ is simply the root mean square value of ϕ for all objects with $\phi \leq 0$ (in the hope that there are very few galaxies with $\phi \leq 0$) computed about the value zero. In other words, we have assumed that the true distribution of ϕ values for stars is symmetric and has a mean of zero, and then computed the *rms* of this distribution. The sum in Equation 5 is performed for all objects within a pre-specified range in calibrated magnitudes, which we chose to be $16.5 \leq r \leq 19.0$. The reason for restricting ourselves to this magnitude range will be explained in § 2.8.

We decided to make $\phi_{\rm cut}$ a multiple of σ. The value of this multiple is somewhat arbitrary, but is kept the same value for each plate. We are now faced with the fundamental compromise of this type of work: if we choose a *large* value of $\phi_{\rm cut}$, we will obtain a catalogue with very few false detections (stars masquerading as galaxies) but which is very incomplete. If we choose a *small* value of $\phi_{\rm cut}$, we will achieve better completeness, but will also include many contaminants, especially since stars outnumber galaxies by a large ratio. The optimal choice depends on the problem to be addressed; in our case (measuring the clustering properties of galaxies), this cut is one which includes about 90% of the galaxies. This is achieved by taking $\phi_{\rm cut} = 3\sigma$.

Below is a list of the adopted values of $\phi_{\rm cut}$ for each field:

Table 2. Values of $\phi_{\rm cut}$

Field #	$\phi_{\rm cut}$	Field #	$\phi_{\rm cut}$
F385	2168	F679	2336
F386	2299	F680	2179
F387	2158	F681	2291
F447	1995	F751	2257
F448	2113	F752	2275
F449	2370	F753	2334
F513	2080	F823	2230
F514	2263	F824	2355
		F825	2380

We point out that because of the way ϕ is defined, it was perhaps not necessary to use a different value of $\phi_{\rm cut}$ for each plate. It might even be argued that a universal value of $\phi_{\rm cut}$ *should* be used. The mean and dispersion of the values of $\phi_{\rm cut}$ in Table 2 are 2240 and 110. Thus, the width of the stellar ridge is being estimated to an accuracy of about 5%. We can estimate the importance of this error on artificial galaxy gradients as follows. We note, as an example, that on Field 449, there are 16192 objects with $\phi > 2240$, and 16390 objects with $\phi > 2140$, thus an error of 100 units of ϕ corresponds to an error of 1.2% in the surface density of galaxies on that plate. Thus, the error in ϕ is comparable to an error of $0.01^{\rm m}$ (see Equation 7).

2.6.4 An Example: F386 cl34

As an instructive example, we present all the relevant data for one of our CCD fields, F386Cl34. The data were obtained on 1989 Apr 7 UT using CCD camera # 8. The exposure times were 1500s in g and 1000s in r. The airmasses were 1.15 and 1.20 in g and r respectively. The photometric transformations for the night are shown on Figures 1 and 2.

Table 4 shows the raw magnitudes g_i and r_i and the reduced magnitudes g and r, as well as the COSMOS magnitudes for all objects common to both CCD frames and the COSMOS catalogue. The columns C_g and C_r are the FOCAS classifications: "g" for galaxy, "s" for star and "sf" for diffuse (uncertain). As noted before, we call an object a galaxy if it is classified as "g" on *both* CCD frames, and a star otherwise. For reference, we have also included the values of ϕ (explained in § 2.6.2) and the raw COSMOS magnitudes, as well as the positions of the objects (from the COSMOS astrometry). The first column is the object number, which corresponds to the numbers shown in Figure 19. This figure is to be compared with Figure 20, which shows a greyscale map of the r CCD image obtained for this cluster.

2.7 Reliability of the Classifier

2.7.1 Comparisons with CCD Frames

It was our hope to select the value of $\phi_{\rm cut}$ independently for each plate by comparing the FOCAS classifications on the CCD data to those of COSMOS. This proved impossible, as there are not enough objects to provide reliable statistics, and the (few) mistakes which FOCAS makes render the problem even more intractable. Fortunately, as already noted and as will be shown below, the method we adopted for determining $\phi_{\rm cut}$ works quite well. For reference, we show the correspondence between the FOCAS and COSMOS classifications in Figure 21. What is shown is a plot of ϕ *vs.* calibrated magnitude for all the objects in the south field. The open squares are the objects classified as galaxies by FOCAS, and the filled symbols are the objects classified as stars by FOCAS. Note that most of the filled symbols lie in the region $\phi < 2240$, while most (but by no means all) open symbols lie in the region $\phi > 2240$. This figure gives a good intuitive understanding of how rapidly the star-galaxy separation deteriorates for faint objects. The filled symbols with very large ϕ values are usually stars with a faint companion which COSMOS is unable to deblend, and hence get classified as galaxies by our ϕ classifier. There is unfortunately nothing which can be done about these objects; the problem is fundamental to the COSMOS deblending software. Recently, the COSMOS group has started using 16 thresholds in the deblending instead of 8, and this should yield some improvements.

2.7.2 Internal Consistency Check

We can use our overlap regions to determine the internal consistency of the ϕ classifier. If we plot the values of ϕ measured on plate 1 *vs.* that measured on plate 2, we obtain a diagram which looks like Figure 22. This plane can be divided into the following regions:

1. $\phi^1 > {\phi_{\rm cut}}^1$ and $\phi^2 > {\phi_{\rm cut}}^2$

   i.e. , both plates agree that the object is a galaxy,

2. $\phi^1 < \phi_{\rm cut}{}^1$ and $\phi^2 < \phi_{\rm cut}{}^2$

i.e. , both plates agree that the object is a star,

3. $\phi^1 < \phi_{\rm cut}{}^1$ and $\phi^2 > \phi_{\rm cut}{}^2$ or $\phi^1 > \phi_{\rm cut}{}^1$ and $\phi^2 < \phi_{\rm cut}{}^2$

i.e. , the plates disagree on the object's classification.

We introduce the following statistic to determine the reliability of the classifications. If the number of objects populating the first two regions above are denoted n_1 and n_2 (shown on Figure 22), and $N_{\rm tot}$ is the total number of objects in the overlap region, then the percentage of agreement between these two plates is

$$\eta = \frac{n_1 + n_2}{N_{\rm tot}} \tag{6}$$

We show the results for all the overlap regions in the following table.

Table 3. Internal Consistency of ϕ Classifier

Overlap	n_1	n_2	$N_{\rm tot}$	η (%)	Overlap	n_1	n_2	$N_{\rm tot}$	η (%)
385:447	258	781	1243	83.5	679:680	283	968	1381	90.5
386:385	453	1425	2122	88.5	679:751	204	912	1243	89.7
386:448	512	1192	1794	94.9	680:681	281	1253	1670	91.8
387:386	725	1581	2420	95.2	680:752	353	1083	1556	92.2
387:449	366	1279	1721	95.5	681:753	407	1270	1773	94.5
448:447	607	1480	2225	93.7	751:752	432	1185	1764	91.6
448:513	749	1608	2590	91.0	751:823	222	843	1160	91.8
449:448	447	1605	2190	93.6	752:753	511	1361	1957	95.6
449:514	560	1472	2135	95.1	752:824	453	1135	1681	94.4
514:513	473	1489	2077	94.4	753:825	475	1112	1668	95.1
					823:824	349	990	1556	86.0
					824:825	709	1098	1907	94.7

The overall repeatability of classification is 92.8%. Of course, this does not tell us that the classifications are *correct*, merely that they are repeatable to high accuracy.

Note the poor results for overlaps with Field 385: this is because the limiting magnitude of this field is not sufficiently deep, and it is after this analysis that we decided to not include this field in our survey.

2.7.3 Visual Tests

We have also performed a comparison of the ϕ classifier with visual inspection of the plates. This was done in the following manner. First, two small regions were selected from the scanned plate, in a region which seemed "typical"†, *i.e.* , not near a bright star, obvious cluster of galaxies, or other feature. We made overlays of these regions (which look just like Figure 19) and then a second overlay consisting only of the images lying in the magnitude range to be considered, namely $16.5 \leq r \leq 19.0$. We then proceeded to inspect the said regions using a stereoscopic magnifier, and noted which objects we thought were stars and which we thought were galaxies.

We must emphasize that this test was performed completely *blind*, in the sense that we had no notion as to what the ϕ values of these objects were at the time of classification. Even then, this test lacks objectivity, in that it is unclear what the learning curve for visual classifications is (*i.e.* , are the classifications done for the tenth plate likely to be more accurate that those for the first plate?), and it is entirely possible that the visual classifications suffer from systematic effects.

As it turned out, there were plates on which I experienced considerably difficulty in deciding upon the classification of the fainter objects, and these are the plates on which the percentage of agreements of my classifications with the ϕ classifier are worst.

The result of these tests are presented in Table 2 of Chapter 4. Apart from the seriously problematic Field 679, the results indicate that *i*) the classifier is very repeatable, and *ii*) the overall accuracy of the classifications is 94%. It is reassuring that this test agrees well with the internal consistency discussed in the previous section.

---

† This is a problem with the CCD fields; being taken near clusters of galaxies, they are highly atypical of the plate as a whole and may thus underestimate the efficiency of the ϕ classifier.

2.8 Range of Usable Magnitudes

The range of magnitudes over which the catalogue can be kept uniform is bounded by two different factors. At the bright end, the main constraint lies in our poor magnitude calibration, which is caused by the paucity of bright objects on the CCD frames. Furthermore, for these bright objects, the cores of the galaxies are saturated on the plate, and their distinct profiles become important, which increases the scatter in the magnitudes even more. Since bright objects are rare, the curve shown on Figure 10 is more uncertain at bright magnitudes than at faint ones. This causes larger plate to plate variations in the zero point of the magnitude scale than are acceptable, so the catalogue must be truncated.

At the faint end, the lack of uniformity comes not from errors in the magnitudes, which are quite good, but rather in varying incompleteness due to each plates different limiting magnitude. This problem also affects the uniformity of the star-galaxy separation. Essentially, the ϕ parameter begins to fail at some fixed distance from the plate's magnitude limit. Hence, in Figure 18, the separation appears to be reasonably good down to $r = 20.0$, and the plate's limiting magnitude is 21.2 (See Table 1, Chapter 4). Many plates do not go as deep as this one, however, so we are limited by the worst plate if we are to keep the catalogue uniform. Field 385 was so bad, in fact, that we chose to eliminate it, since its inclusion would have forced us to truncate the catalogue at somewhere around 18^{m}, which is unacceptably shallow.

The effects of these non-uniformities can be seen in Figures 23, 24, 25, 26, and 27, which show dot plots for the north catalogue for objects lying in the magnitude interval (0.0,16.0), (16.0,17.0) ... (19.0,20.0). The plate boundaries are quite apparent in the first and last of these figures. The black squares are regions which have been deleted due to bright stars, satellite trails, emulsion defects *etc.*

These constraints, then, have led us to use only objects lying in the magnitude range $16.5 \leq r \leq 19.0$. Figures 28 and 29 show dot plots for all the galaxies in this magnitude range for the north and south fields. For clarity, the deleted regions have not been plotted, which is why some blank regions can be found.

This magnitude range is unfortunately rather narrow. The only remedy to this problem is to get better plates. Note, for example, how for the four bottom left plates in the north field (which are all quite deep), none of the boundaries can be seen between these plates on Figure 27.

2.9 Pitfalls of the Overlap Regions

Finally, we would like to discuss a topic which is particularly important in light of the method being used by other groups in the preparation of large galaxy catalogues. I am referring of course to the use of the overlap region to internally calibrate a large mosaic of plates, and then using sparse CCD photometry to establish the overall zero point of the magnitude scale. We wish to demonstrate in this section that this is a potentially dangerous method to use. We note that our catalogue is the *only* one to date with sufficient CCD calibration sequences to check the validity of this method in a direct manner.

In our survey, the magnitude scale of each plate was *independently* calibrated using the CCD data. In this method, it is inevitable that some small errors in the zero point of the calibration scales of each plate remain. In fact, we even *know* how large these errors should be: the *rms* of the observed zero points for each plate about the mean should be $0.06^{\rm m}$, as calculated in § 2.5.1.

The effect of a small magnitude scale difference δm between two plates will be to cause a difference in surface density of galaxies †

$$\frac{\delta\sigma}{\sigma} \approx \frac{\frac{d}{dm}(10^{0.45m})\,\delta m}{10^{(0.45m)}} = \ln(10)\ 0.45\ \delta m \qquad (7)$$

Thus a difference in magnitudes of $0.05^{\rm m}$ yields an increase in surface density of objects of 5%! The effects of such spurious density variations on measurements of the galaxy-galaxy correlation function $w(\theta)$ are discussed by Geller *et al.* (1984).

However, we still have information which has not been used; that contained in the overlap regions. Figure 30 shows a plot of calibrated magnitudes for galaxies in the overlap region of fields 448 and 449. The straight line is the $y = x$ line about which the points would scatter if the calibration on each plate was perfect. We should be able to use this information to "tie together" the various magnitude scales, and thus further reduce the artificial gradients in galaxy density across our fields.

† We will show in Chapter 4 that the slope of the galaxy number counts is 0.45.

We have attempted to determine corrections to be applied to the zero point of the magnitude scales for each plate as follows. Let $F_{ij}(m)$ be the function which transforms magnitudes from plate i to plate j. For simplicity, we will assume that the F_{ij} are linear functions, which we derive by fitting a straight line to the points in Figure 30. Since these functions will be used only on a very narrow range of magnitudes, the assumption of linearity is good. It is important to note that these functions must be *antisymmetric*: if F_{12} says that plate 2 has an offset of $+0.1^{\rm m}$ to plate 1, then F_{21} had better yield an offset of $-0.1^{\rm m}$ for plate 2 *vs.* plate 1. We accomplished this by minimizing the perpendicular distance to the best line from each point. This is necessary, as there is scatter in each of the two axes in Figure 30, so a simple least-squares fit using the y-deviation systematically underestimates the slope of the best fitting line. Because we have no better guess, we assumed that the measurement errors were the same on each plate (*i.e.* , the error boxes are square, not rectangular.) The large number of galaxies present in the overlap region allows us to determine the $F_{ij}(m)$'s to high accuracy; the uncertainty in the fit at $19^{\rm m}$ is approximately $0.03^{\rm m}$. Once we have $F_{ij}(m)$, it is a simple matter to invert the function and obtain $F_{ji}(m)$.

We now consider the small difference $\delta_{ij} = m_j - F_{ij}(m_i)$ where m_i and m_j are the proposed magnitude cutoffs for plates i and j. Physically, the δ_{ij}'s are the difference between the adopted magnitude cutoff on plate j and what plate i predicts this cutoff should be (using its own cutoff value). We propose the following iterative scheme to get a self-consistent solution for the m_i's which minimizes $\epsilon = \sum_{\text{all overlaps}} \delta_{ij}^2$.

Let the plates be numbered $1, 2 \ldots 9$, with plate 5 being in the center, plates 1,3,7 and 9 in the corners (the orientation is irrelevant). We first choose a cutoff magnitude (say $m_{cut} = 19.0$), set $m_j = m_{cut}\ \forall\ j$ and calculate the quantities $F_{ij}(m_i)$. Then we replace the values of the m_i's in the following manner:

$$m_5 \rightarrow \frac{1}{4}(F_{25}(m_2) + F_{45}(m_4) + F_{65}(m_6) + F_{85}(m_8)) \tag{8}$$

for the central plate,

$$m_2 \rightarrow \frac{1}{3}(F_{12}(m_1) + F_{32}(m_3) + F_{52}(m_5)) \tag{9}$$

and so on for the other three plates on the edges, and

$$m_1 \rightarrow \frac{1}{2}(F_{41}(m_4) + F_{21}(m_2))/2 \tag{10}$$

and similarly for the other three plates in the corners.

After each iteration we recalculate $\epsilon = \sum \delta_{ij}^2$ and stop iterating when it no longer diminishes. We then have a set of corrections $\alpha(j) = m_j - m_{cut}$ which we can apply to each plate before joining them together in the final catalogue. Unfortunately, in our north field, two of the plates could not be used (fields 385 and 512) corresponding to plates 3 and 9 in this scheme; we have therefore simply eliminated the appropriate equations from the iterative scheme. However, field 447 (plate 6) then has only one overlap region (with plate 5) so this scheme resorts to simply fixing the magnitude m_6 so that $m_6 = F_{56}(m_5)$.

We show in Figures 31, 32, 33, 34 and 35 the progress of this method. In these figures, the central (large) numbers are the magnitude limits of each plate (the m_i's) and the small numbers are the values predicted for the m_i's by the neighbouring plate (the $F_{ij}(m_i)$'s). In practice, this scheme converges very quickly; the solutions became stable after only 4 iterations.

The method is spectacularly successful at finding a solution whose error value is very small; the values for ϵ in Figures 31 through 35 are 10.6, 8.9, 5.7, 5.2 and 5.1, after which ϵ stops improving. In the north, ϵ drops from 5.0 to 0.8, but there are four fewer overlap regions because of the missing fields, so we do not take this as indication that the north field is better calibrated than the south.

However, this scheme requires that we make adjustments to the zero points of the plates which are several times larger than our formal uncertainty in the calibration procedure of 0.06^{m}(derived in § 2.5.1). Note for example, the difference in Figure 35 between plates 1 and 9, they differ by 0.4^{m}! And of course, the fact that the initial value of ϵ is much larger than 0.6^{m} is worrisome. The two methods are therefore irreconcilable, and we are forced to choose between them.

We have chosen to interpret this discrepancy as being due to an intrinsic non-uniformity in the magnitude scale of the plate near the edges, and not due in any way to errors in the CCD photometry. The edges of the plates are vignetted, as shown on Figure 4, and so we must conclude that the magnitudes in the vignetted regions are not being properly compensated for by COSMOS. In fact, it will be seen from the analysis of the star counts in the appendix to Chapter 3 that this suspicion is well-founded, and that there is indeed a paucity of objects near the edges of the plates. We therefore conclude that serious errors can propagate through a network of plates whose magnitudes are calibrated using only the overlap regions. In this catalogue, we have *not* included the type of corrections discussed in this section.

Table 4 Data for F386 Cl34

#	g_i	r_i	COSMAG	C_g	C_r	ϕ	r	g	$g-r$	α_{1950}			δ_{1950}		
24	21.30	19.92	20.03	g	g	−447	20.32	22.59	2.28	14	53	54.1	33	10	25
25	16.83	16.55	16.97	s	s	−1244	16.96	17.96	1.01	14	54	1.4	33	10	35
26	20.53	20.02	20.11	g	g	138	20.42	21.70	1.27	14	53	59.5	33	10	41
27	18.30	18.93	18.62	s	s	3638	19.34	19.30	−0.04	14	53	55.1	33	12	8
28	19.56	19.74	19.71	g	g	1265	20.15	20.63	0.48	14	53	54.2	33	12	13
29	20.35	19.95	19.54	s	s	1675	20.35	21.50	1.15	14	54	2.5	33	12	14
30	20.24	20.09	19.40	g	g	2464	20.50	21.35	0.86	14	54	0.1	33	12	16
31	18.44	18.16	18.00	s	s	679	18.57	19.57	1.01	14	53	59.7	33	13	36
32	19.12	19.18	18.84	s	s	688	19.59	20.20	0.62	14	53	56.1	33	13	52
33	19.57	20.27	19.75	g	g	3167	20.68	20.56	−0.12	14	54	2.3	33	14	28
35	14.88	15.18	15.74	s	s	270	15.59	15.93	0.34	14	53	55.1	33	14	57
36	16.64	17.10	17.22	s	s	404	17.51	17.67	0.15	14	53	53.9	33	15	7
37	20.27	20.27	20.98	g	g	−641	20.68	21.36	0.68	14	54	2.5	33	15	29
38	16.62	16.59	15.95	g	g	17912	17.00	17.72	0.72	14	53	55.4	33	15	51
39	19.17	19.21	18.55	g	g	5024	19.62	20.26	0.64	14	53	55.2	33	16	2
40	18.48	18.40	17.75	g	g	7818	18.81	19.58	0.78	14	54	2.6	33	15	53
42	20.43	20.33	20.66	g	g	−451	20.74	21.54	0.80	14	54	1.3	33	16	20
43	19.71	19.60	19.24	g	g	−333	20.01	20.82	0.81	14	53	58.5	33	16	24
44	19.01	20.34	20.58	g	s	−714	20.76	19.91	−0.85	14	53	58.4	33	16	31
45	19.49	19.41	19.44	g	g	5456	19.82	20.59	0.78	14	54	2.3	33	16	37
46	17.59	17.71	17.26	g	g	4721	18.12	18.67	0.55	14	53	56.6	33	16	53
47	20.90	20.67	20.36	g	g	−68	21.08	22.03	0.95	14	54	2.7	33	16	54
48	16.65	16.98	17.18	s	s	−976	17.39	17.69	0.30	14	53	53.7	33	17	6
49	18.36	18.64	18.14	g	g	15709	19.05	19.41	0.36	14	53	59.3	33	17	5
50	17.42	17.46	17.21	g	g	1612	17.87	18.51	0.64	14	53	55.4	33	17	6
51	18.25	18.28	17.90	g	g	4167	18.69	19.34	0.65	14	54	2.7	33	17	4
52	21.06	20.35	20.90	g	sf	705	20.75	22.26	1.50	14	53	57.4	33	17	17
53	20.20	20.19	19.93	g	g	2186	20.60	21.29	0.70	14	53	54.4	33	17	35
54	19.35	19.35	19.07	g	g	729	19.76	20.44	0.68	14	53	53.2	33	17	38
55	18.59	18.25	17.99	g	g	15784	18.66	19.73	1.08	14	53	54.6	33	17	54
56	20.47	20.99	21.02	s	g	−366	21.40	21.49	0.08	14	53	59.5	33	18	4
57	21.15	20.20	20.11	sf	g	463	20.60	22.38	1.78	14	54	2.2	33	18	4
67	19.40	19.86	20.31	g	g	1152	20.27	20.43	0.15	14	53	50.0	33	10	17
68	16.50	16.82	17.11	s	s	81	17.23	17.55	0.32	14	53	48.7	33	10	27
69	18.21	18.73	18.49	s	s	268	19.14	19.23	0.08	14	53	49.3	33	10	38
70	18.97	19.25	19.17	g	g	3532	19.66	20.02	0.36	14	53	40.8	33	11	0
71	21.73	20.26	20.57	g	g	3827	20.66	23.04	2.38	14	53	51.4	33	11	9
72	16.90	16.96	16.43	g	g	10526	17.37	17.98	0.62	14	53	45.8	33	11	39
73	15.31	15.43	16.09	s	s	221	15.84	16.39	0.55	14	53	42.6	33	11	57
74	17.51	17.43	16.98	g	g	5460	17.84	18.61	0.78	14	53	49.4	33	12	16
75	16.08	16.30	16.52	g	g	−354	16.71	17.14	0.43	14	53	51.8	33	12	15

Table 4 Data for F386 Cl34 (continued)

#	g_i	r_i	COSMAG	C_g	C_r	ϕ	r	g	$g-r$	α_{1950}	δ_{1950}
76	20.91	20.54	20.83	g	g	−9	20.95	22.06	1.11	14 53 50.2	33 13 31
77	17.07	17.48	17.36	s	sf	504	17.89	18.10	0.21	14 53 48.0	33 14 4
78	20.99	20.80	20.95	g	g	1069	21.21	22.11	0.90	14 53 41.3	33 14 6
79	19.54	19.12	18.73	s	g	1714	19.52	20.69	1.17	14 53 42.9	33 14 7
80	20.15	20.37	20.52	g	g	1452	20.78	21.21	0.43	14 53 50.9	33 14 34
81	17.52	17.47	17.10	g	g	2702	17.88	18.62	0.74	14 53 46.9	33 14 36
82	18.71	19.04	18.27	g	g	10766	19.45	19.75	0.30	14 53 45.9	33 14 40
83	17.13	17.11	16.66	g	g	6684	17.52	18.23	0.71	14 53 42.1	33 14 42
84	18.56	18.14	17.82	s	sf	273	18.54	19.71	1.17	14 53 49.3	33 14 58
85	20.30	20.39	20.00	g	g	−307	20.80	21.38	0.58	14 53 49.1	33 15 7
86	19.29	19.60	19.32	sf	s	−490	20.01	20.34	0.33	14 53 46.8	33 15 12
87	19.95	19.92	19.98	g	g	2076	20.33	21.05	0.72	14 53 41.7	33 15 21
88	21.24	21.35	20.86	g	g	727	21.76	22.32	0.56	14 53 47.3	33 15 21
89	16.98	16.87	16.35	g	g	7114	17.28	18.09	0.81	14 53 45.3	33 15 22
90	18.50	19.05	18.12	g	g	29546	19.46	19.51	0.05	14 53 46.8	33 15 24
91	16.71	17.21	17.16	s	s	941	17.62	17.73	0.11	14 53 47.4	33 15 29
92	19.88	19.91	19.84	g	g	3419	20.32	20.97	0.65	14 53 42.8	33 16 6
93	19.01	18.44	18.16	s	s	−792	18.84	20.19	1.34	14 53 47.2	33 16 8
94	18.73	18.62	18.02	g	g	6057	19.03	19.84	0.81	14 53 43.4	33 16 31
95	19.30	19.23	18.86	s	s	422	19.64	20.40	0.76	14 53 48.4	33 16 31
96	19.75	19.80	19.60	g	g	1537	20.21	20.84	0.63	14 53 50.2	33 16 53
97	19.34	19.39	19.33	g	g	−325	19.80	20.43	0.63	14 53 46.5	33 17 7
98	19.56	19.15	18.88	s	s	−14	19.55	20.71	1.16	14 53 42.5	33 17 18
99	15.84	15.79	15.35	g	g	13160	16.20	16.94	0.74	14 53 50.0	33 17 51
100	15.94	16.04	15.50	g	g	15204	16.45	17.02	0.57	14 53 45.8	33 18 0
115	17.50	17.37	16.92	g	g	10398	17.78	18.61	0.83	14 53 34.2	33 10 27
116	18.83	18.76	18.35	g	g	14902	19.17	19.93	0.76	14 53 32.9	33 10 42
117	16.28	16.74	17.11	s	s	737	17.15	17.31	0.15	14 53 30.7	33 10 56
118	21.26	20.42	20.20	g	g	3397	20.82	22.47	1.65	14 53 40.2	33 11 15
119	18.03	17.49	17.54	s	s	261	17.89	19.20	1.31	14 53 30.9	33 11 20
120	20.43	19.62	19.72	g	g	2381	20.02	21.64	1.62	14 53 38.4	33 11 25
121	19.47	19.02	18.62	s	s	−415	19.42	20.63	1.20	14 53 40.4	33 11 39
122	16.90	17.49	17.42	s	s	−119	17.90	17.91	0.00	14 53 34.6	33 12 10
123	17.77	17.98	17.34	g	g	7385	18.39	18.83	0.44	14 53 38.2	33 12 28
124	18.20	18.36	17.91	g	g	5371	18.77	19.27	0.50	14 53 31.1	33 12 31
125	19.85	19.99	20.21	s	sf	−1155	20.40	20.92	0.52	14 53 31.6	33 12 44
126	18.83	18.40	18.13	s	s	−79	18.80	19.98	1.18	14 53 35.8	33 13 7
127	20.44	20.49	21.26	g	g	5939	20.90	21.53	0.63	14 53 35.2	33 13 33
128	20.52	20.32	20.35	s	s	−306	20.73	21.64	0.91	14 53 32.2	33 13 53
130	18.34	18.41	17.91	g	g	4166	18.82	19.42	0.60	14 53 34.4	33 14 21
131	14.81	15.27	15.83	s	s	188	15.68	15.84	0.15	14 53 30.7	33 14 53

Table 4 Data for F386 Cl34 (continued)

#	g_i	r_i	COSMAG	C_g	C_r	ϕ	r	g	$g-r$	α_{1950}	δ_{1950}
132	19.61	19.65	19.46	g	g	722	20.06	20.70	0.64	14 53 30.5	33 15 6
133	19.53	19.21	18.66	g	g	19097	19.62	20.67	1.05	14 53 35.9	33 15 10
134	17.63	17.67	16.94	g	g	14408	18.08	18.72	0.64	14 53 38.7	33 15 34
135	17.97	17.97	17.63	s	s	437	18.38	19.06	0.68	14 53 39.0	33 15 43
136	20.02	20.27	20.42	g	g	561	20.68	21.08	0.40	14 53 33.0	33 15 46
137	20.23	20.19	20.18	g	g	504	20.60	21.33	0.73	14 53 38.7	33 15 58
138	17.72	18.32	17.93	s	s	244	18.73	18.73	-0.01	14 53 41.2	33 16 0
139	15.93	16.38	16.25	s	s	-345	16.79	16.96	0.17	14 53 35.5	33 16 5
140	18.24	18.20	17.55	g	g	13346	18.61	19.34	0.73	14 53 41.3	33 16 10
143	18.50	18.84	18.31	g	g	8578	19.25	19.54	0.29	14 53 31.8	33 16 17
144	16.08	16.25	15.73	g	g	6430	16.66	17.15	0.49	14 53 39.0	33 16 16
145	20.67	20.16	20.22	s	s	-407	20.56	21.84	1.27	14 53 31.6	33 17 10
161	17.57	17.89	17.66	s	g	584	18.30	18.62	0.32	14 53 28.0	33 11 22
162	20.81	20.59	21.14	g	g	1317	21.00	21.93	0.94	14 53 28.5	33 11 43
164	20.15	19.81	19.70	g	s	-77	20.22	21.29	1.08	14 53 26.0	33 13 2
168	19.63	20.32	19.82	s	s	1007	20.73	20.62	-0.11	14 53 30.0	33 13 40
170	20.32	19.78	19.48	g	g	489	20.18	21.49	1.31	14 53 29.0	33 14 16
172	19.89	20.03	19.89	s	s	23	20.44	20.96	0.52	14 53 30.1	33 14 32
176	19.96	19.61	19.48	g	g	1354	20.02	21.10	1.09	14 53 27.1	33 16 14
177	17.97	17.87	17.83	s	s	-329	18.28	19.08	0.80	14 53 29.8	33 16 41
178	19.43	18.88	18.71	g	sf	541	19.28	20.60	1.32	14 53 26.2	33 16 53
179	17.83	17.86	17.52	g	g	5149	18.27	18.92	0.65	14 53 28.0	33 16 54
181	17.83	17.83	17.77	s	s	-605	18.24	18.92	0.68	14 53 26.9	33 17 0
183	16.51	16.89	17.18	s	s	-332	17.30	17.55	0.25	14 53 29.3	33 17 34
186	16.68	16.81	17.15	s	s	-689	17.22	17.75	0.53	14 53 28.1	33 17 57

2.10 Figure Captions

Figure 1. The photometric solution obtained in the g band from observation of standard stars for the night of 1989 April 7 UT.

Figure 2. Same as Figure 1, but for the r band.

Figure 3. Histograms of the number of galaxies found in each magnitude bin, for four clusters of galaxies. The left panels are in the r band, and the right for the g band. The exposures are typical. Galaxies are easily detected down to $22^{\rm m}$ in r and $23^{\rm m}$ in g.

Figure 4. A contour plot of the sky values (in units of log intensity) on Field 449 which are subtracted from the images to obtain the COSMOS magnitudes. The image is 5.3° on a side, and the contours range from 3.96 to 4.77 by steps of 0.01. Notice the gradients around bright stars, and the vignetting near the edges and corners of the plates.

Figure 5. Objects found by COSMOS in the vicinity of a bright star. Note the diffraction peaks, and the multitude of faint images deblended from the halo of the star; none of those are real. Such objects were removed from the catalogue manually.

Figure 6. COSMOS catalogue near a rich cluster on Field 449. The detected images are represented by ellipses with parameters *a_i, b_i* and *theta_i* (see Table 1). The parent images are superposed on their deblended counterparts as the very large ellipses. These are shown for illustration only; normally all the parent images are rejected.

Figure 7. The raw (uncalibrated) COSMOS magnitudes for objects common to both Fields 448 and 449 plotted against one another. Note the large zero point offset (which is removed by our calibration procedure) and the different behaviours of stars and galaxies, especially apparent for COSMAG < 1600.

Figure 8. The location of the CCD calibration sequences on the photographic plates are shown for the north field. The solid lines denote the boundaries of the survey, and the dashed lines the location where the plates were pasted together. The locations of the plate boundaries is approximate, and the size of the CCD sequences is not to scale, but greatly enlarged for clarity.

Figure 9. Same as Figure 8, but for the south field.

Figure 10. The calibration curve used for Field 752. The calibrated FOCAS magnitudes are plotted against the raw COSMOS magnitudes for four clusters of galaxies. We show each cluster using a different symbols to check for non-uniformities in the magnitude scale across the plate. Note that there no systematic differences are seen between symbols.

Figure 11. The difference in the calibrated magnitudes for objects common to both Fields 448 and 449 against their averages. In a perfect calibration, the points will scatter around the line $y = 0$.

Figure 12. A histogram of the number of objects as a function of the calibrated magnitude for all objects on Field 449. The limiting magnitude is defined to be where the histogram turns over, in this case at $r = 21.2$.

Figure 13. The area of objects (in pixels) *vs.* the raw COSMOS magnitude. Note the sharp stellar ridge, and the outliers, which are galaxies. These galaxies become indistinguishable from stars near COSMAG = 1700 (*c.f.* Figure 14).

Figure 15. A histogram of the $\mathcal{S}$ values in the range $18.1 <$ COSMAG < 18.3. The values of the mode and the width of the distribution are also indicated. Figures 14, 15, 16 and 17 are presented out of order so that they can be overlayed properly, to illustrate the fitting procedure to the stellar ridge.

Figure 17. An example of the smoothing spline fitting algorithm used to fit a function to the $\mathcal{S}$ parameter and its width. Note how, unlike a regular spline, the points are not constrained to pass through each data point.

Figure 16. The values obtained from the distribution shown on Figure 14 plotted versus COSMAG. The widths are obtained from the difference of the modal value and the 5^{th} percentile. Note that these trace the distribution of $\mathcal{S}$ values shown on Figure 14 very well.

Figure 14. $\mathcal{S}$ is plotted *vs.* raw COSMOS magnitude. Here the separation is reasonably clean at least to COSMAG = 1850 (*c.f.* Figure areavsmag).

Figure 18. A plot of ϕ *vs.* the calibrated r magnitude for Field 449. The stellar ridge is clearly defined about $\phi = 0$. The separation is clean out to approximately 20^{m} on this plate; fainter than this, all objects are merged into the stellar ridge.

Figure 19. An overlay of the cluster F386cl34, indicating the object numbers corresponding to those in Table 4.

Figure 20. A greyscale of the r CCD image obtained for the cluster F386cl34.

Figure 21. The correspondence between the ϕ classifier and FOCAS is shown. The ϕ value for all the objects found on the CCD frames used to calibrate the south fields are plotted against their calibrated r magnitudes. The filled symbols are objects classified as stars by FOCAS, and the open symbols thos classified as galaxies. Note how most FOCAS stars lie on the stellar ridge defined by $\phi = 0$, but there are several outliers. These are usually merged objects on which the COSMOS deblending software failed, or whose parameters were not well reproduced, and so end up with a large ϕ value. For $r > 19$, several objects classified as galaxies by FOCAS have values of ϕ small enough that they get classified as stars, and so the galaxy catalogue becomes incomplete.

Figure 22. The ϕ values measured for objects common to both Fields 448 and 449. The dense, circular feature is the stellar ridge. The tail of this "comet" represents the galaxies.

Figure 23. A dot plot of galaxies brighter than $r = 16.0$ for the north field. The black squares are deleted regions around bright stars and satellite trails. Note the lack of uniformity; for example, the apparent lack of galaxies on the top right and bottom left plates.

Figure 24. Same as Figure 23, but for galaxies with magnitudes $16.0 \leq r \leq 17.0$.

Figure 25. Same as Figure 23, but for galaxies with magnitudes $17.0 \leq r \leq 18.0$.

Figure 26. Same as Figure 23, but for galaxies with magnitudes $18.0 \leq r \leq 19.0$.

Figure 27. Same as Figure 23, but for galaxies with magnitudes $19.0 \leq r \leq 20.0$.

Figure 28. A dot plot of the north catalogue, for all objects in the nominal magnitude range $16.5 \leq r \leq 19.0$. The deleted regions have not been plotted for clarity.

Figure 29. Same as Figure 28, but for the south field.

Figure 30. Same as for Figure 7, but this time for the calibrated magnitudes. The line shown is not a fit: it is the $y = x$ line about which points would scatter if the calibration was perfect.

Figure 31. The progress of the iterative algorithm used to find corrections to the magnitude scales of the south plates is shown. Each panel represents one plate. The large numbers are the current adopted magnitude cuts on each plate (in magnitudes multiplied by 100), and the small numbers are the values predicted from the neighbouring plate, using its own present cutoff value. Thus, in this figure, if the top left plate is plate #1 and the top middle plate is plate #2, we obtain $F_{21}(1900) = 1871.9$. All cutoff magnitudes are initially set at 1900 and then are evolved according to the procedure described in § 2.2.9.

Figure 32. Same as Figure 31, after 1 iteration.

Figure 33. Same as Figure 31, after 2 iteration.

Figure 34. Same as Figure 31, after 3 iteration.

Figure 35. Same as Figure 31, after 4 iteration.

Figure 1.

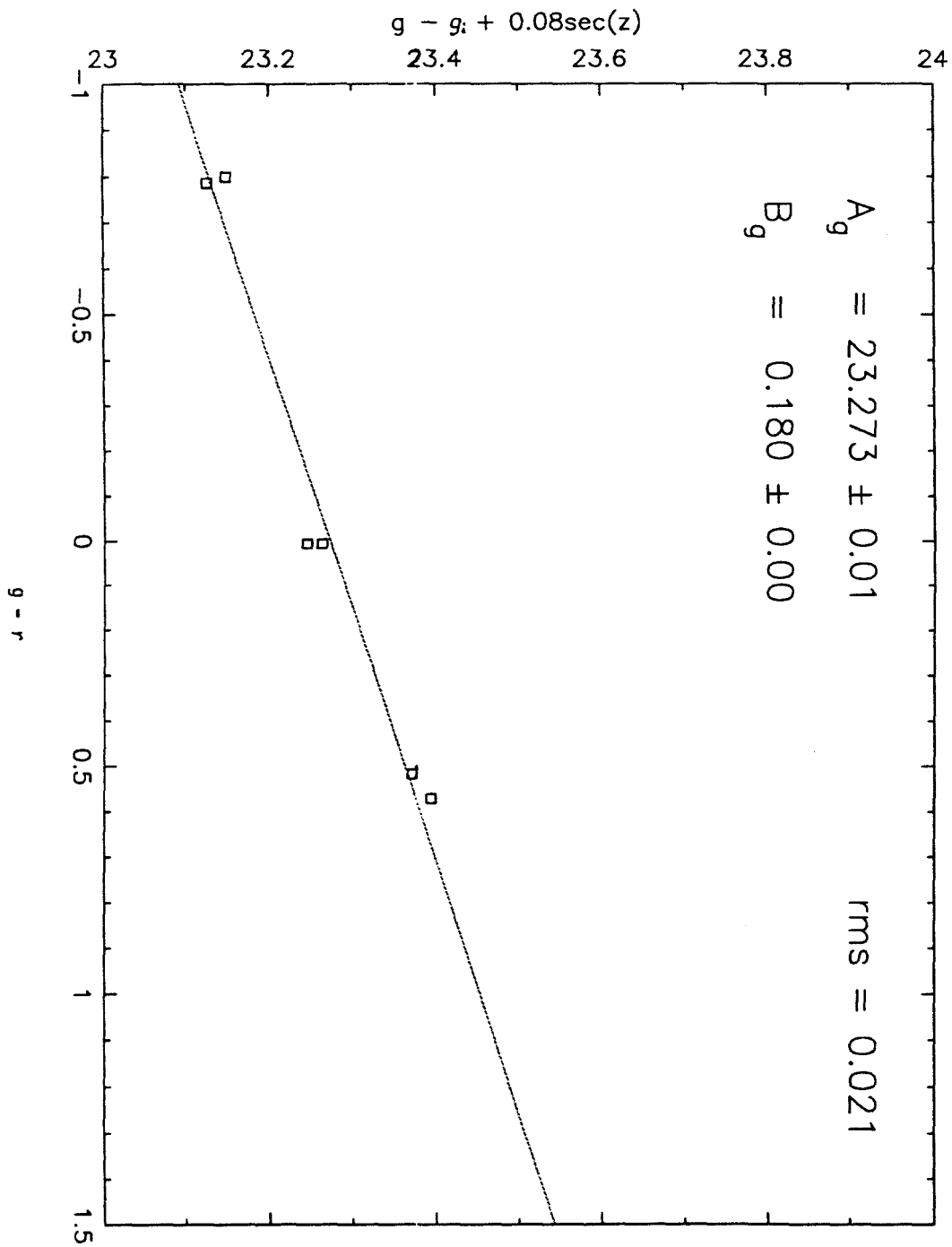

g − g_i + 0.08sec(z)
23
23.2
23.4
23.6
23.8
24
−1
−0.5
0
0.5
1
1.5
g − r
A_g = 23.273 ± 0.01
B_g = 0.180 ± 0.00
rms = 0.021
April 07 1989 UT

Figure 2.

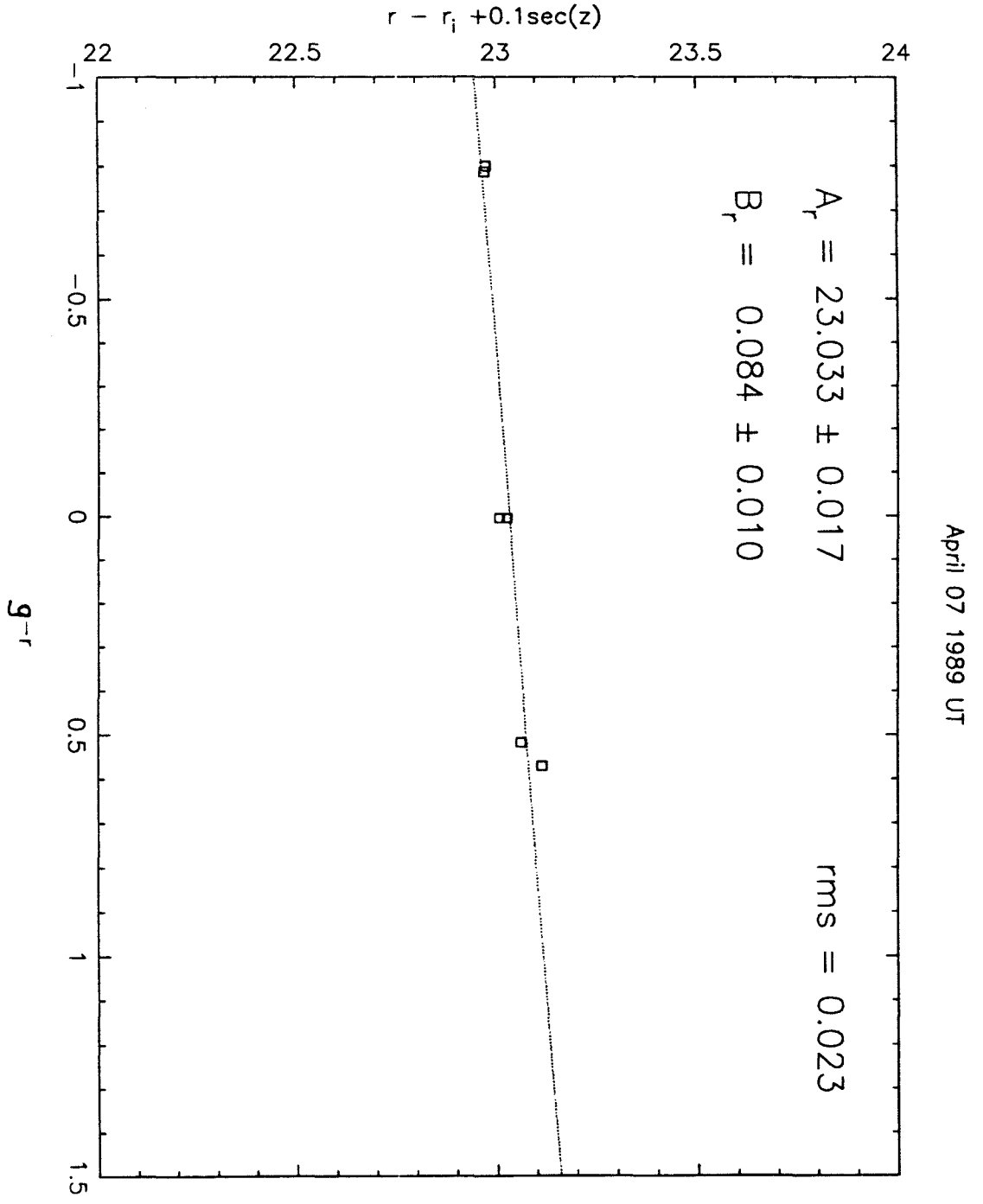

r − r_i +0.1sec(z)
22
22.5
23
23.5
24
−1
−0.5
0
0.5
1
1.5
g−r
A_r = 23.033 ± 0.017
B_r = 0.084 ± 0.010
rms = 0.023
April 07 1989 UT

Figure 3.

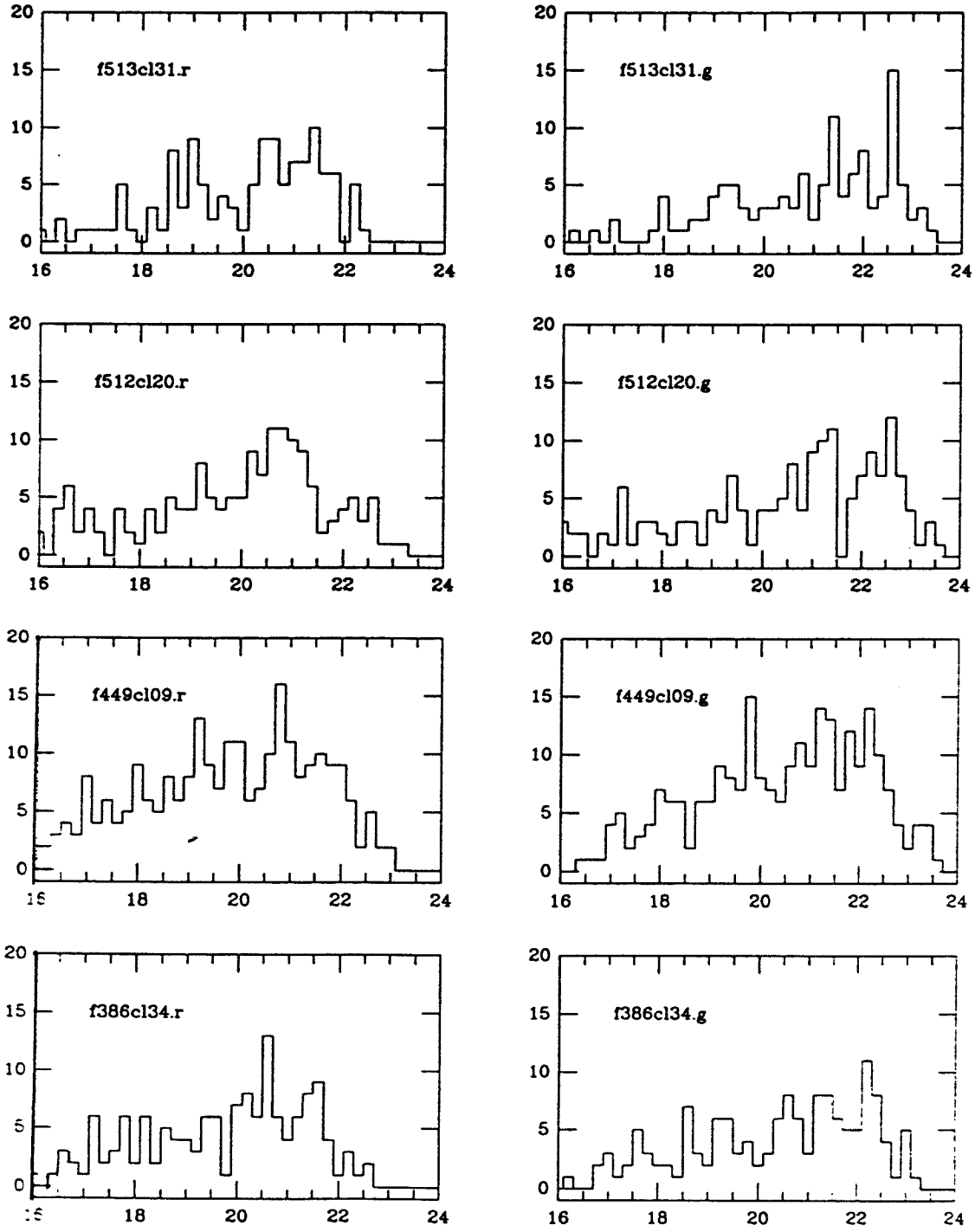

f513cl31.r
f513cl31.g
f512cl20.r
f512cl20.g
f449cl09.r
f449cl09.g
f386cl34.r
f386cl34.g
0
5
10
15
20
16
18
20
22
24

Figure 4.

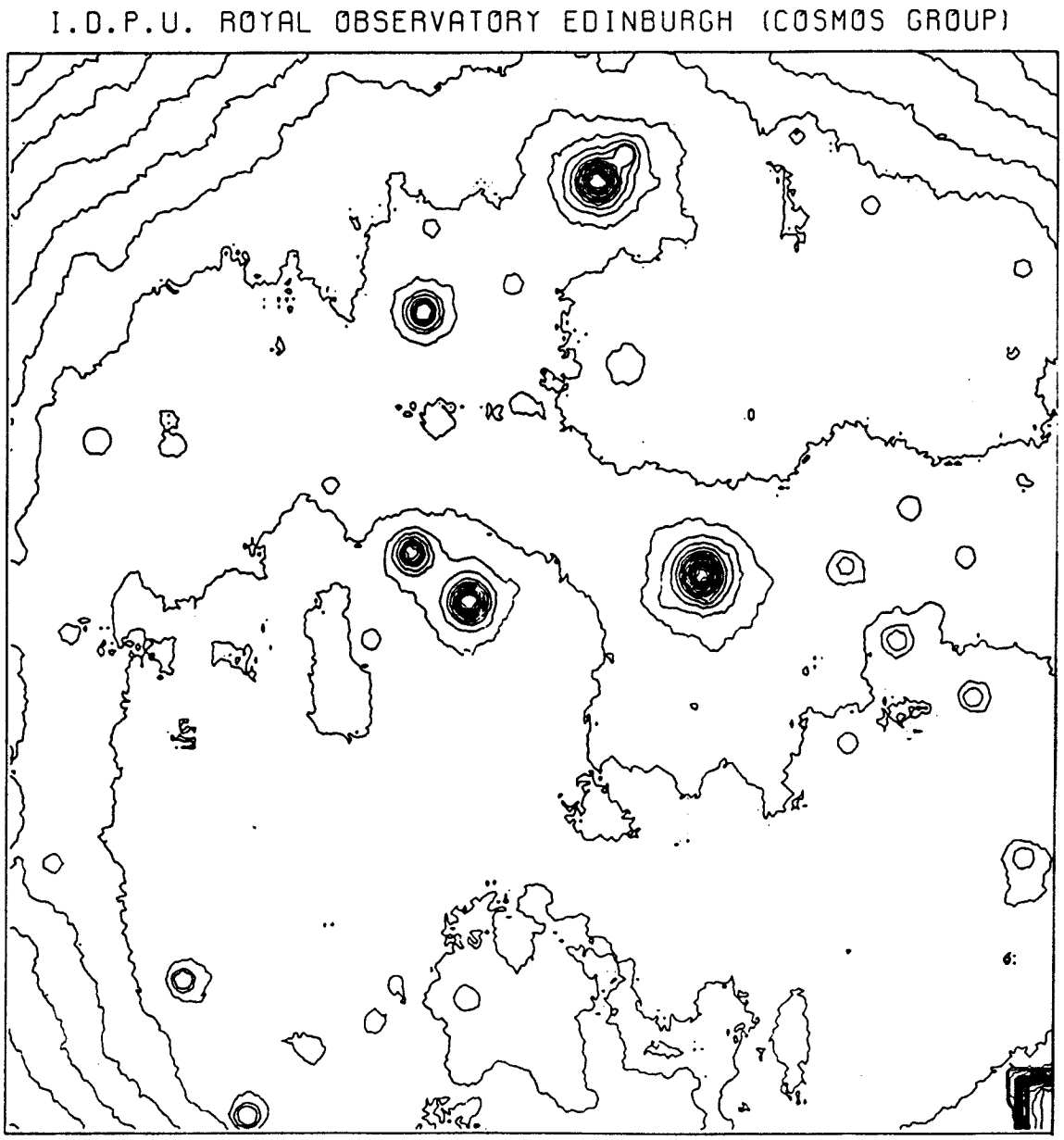


Figure 5.

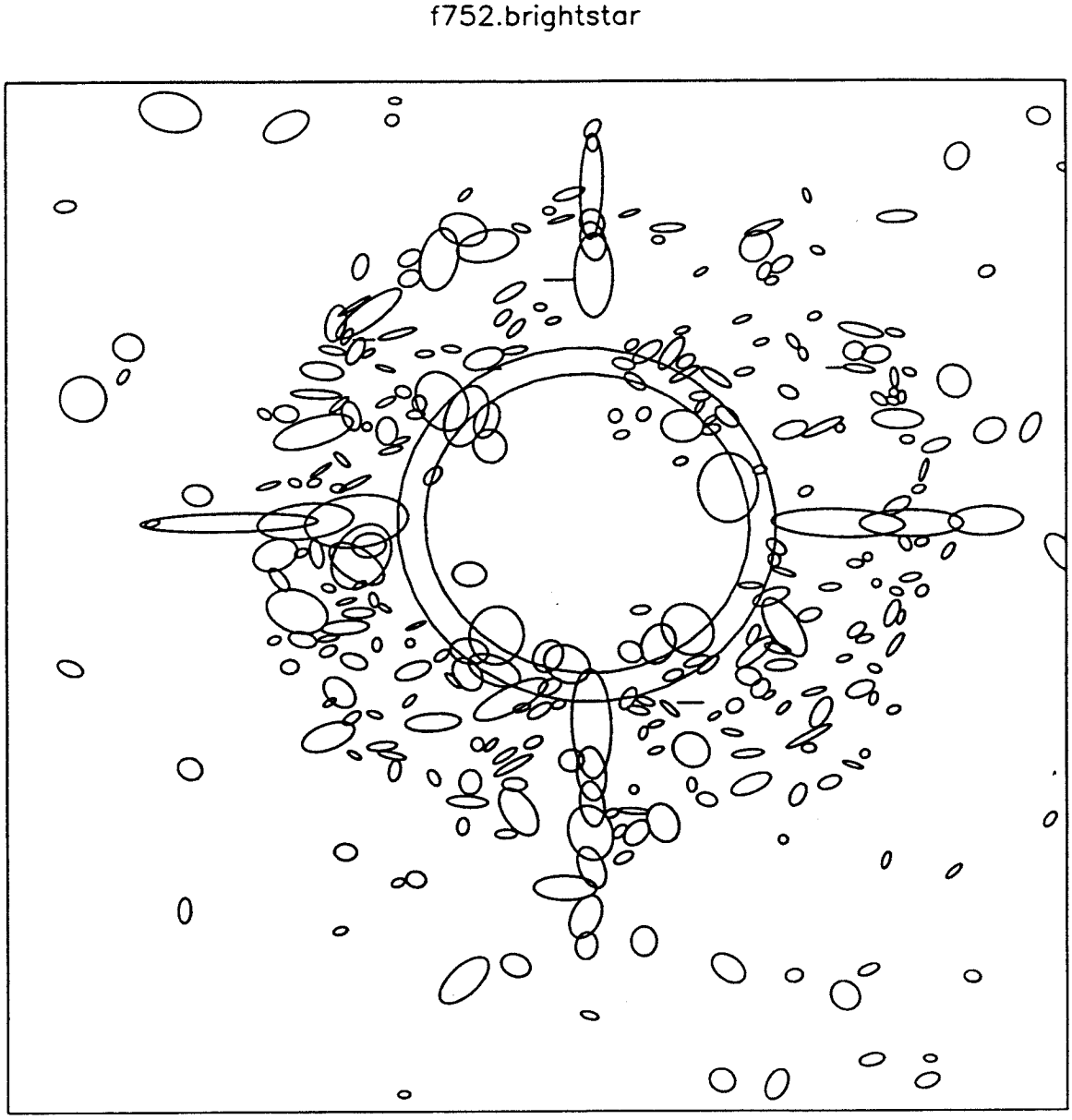


Figure 6.

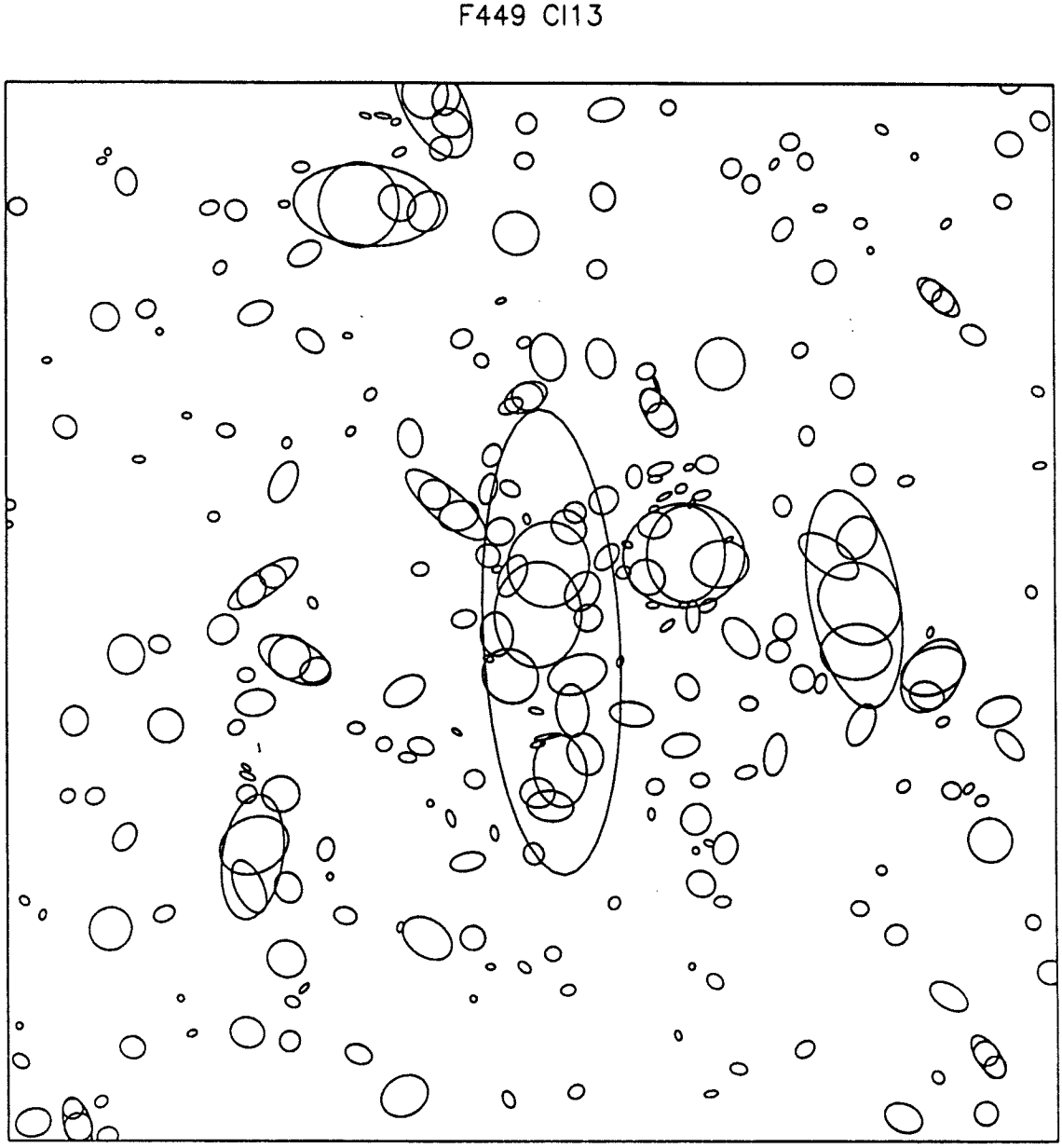


Figure 7.

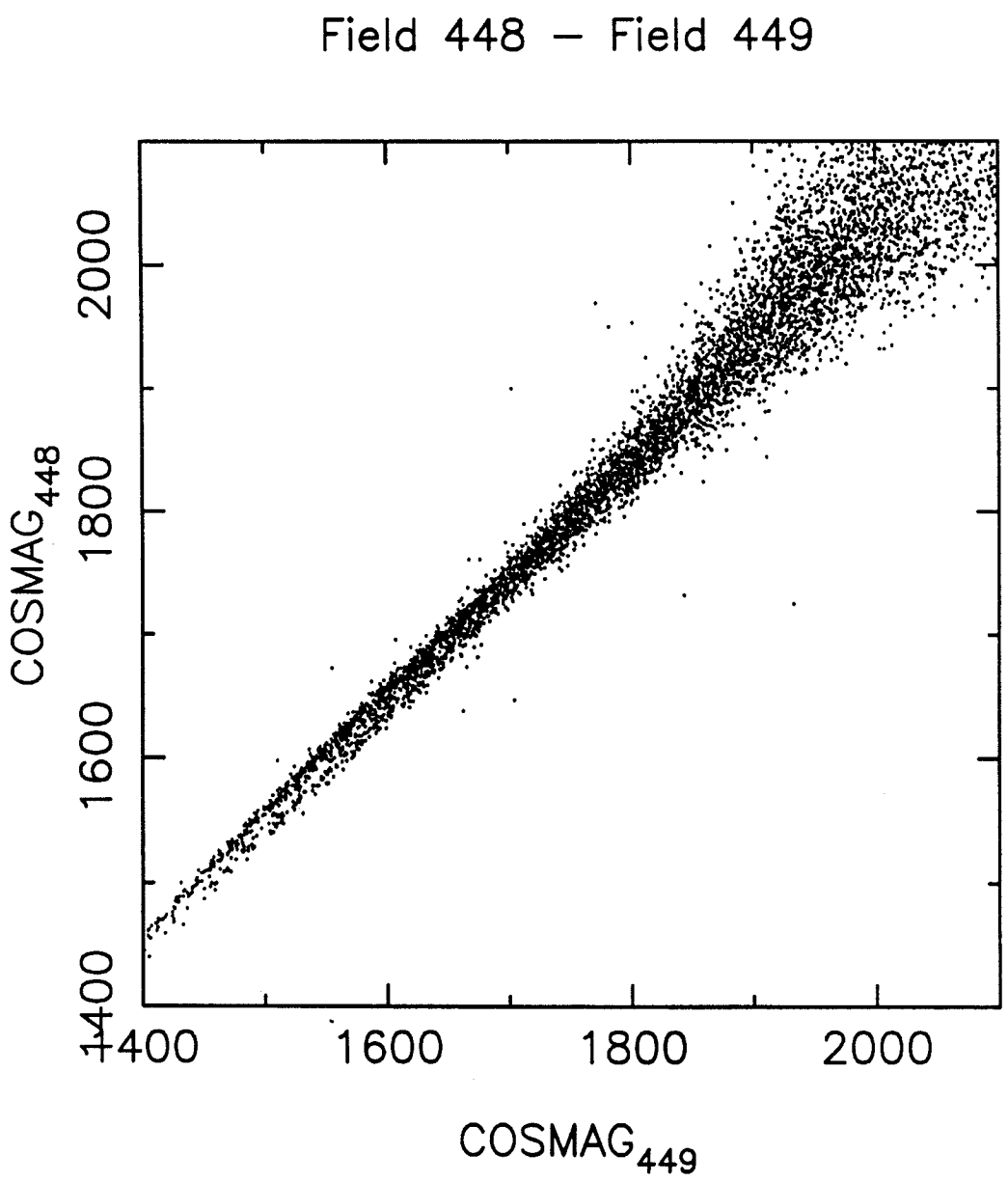

Field 448 – Field 449
COSMAG448
COSMAG449
1400
1600
1800
2000

Figure 8.

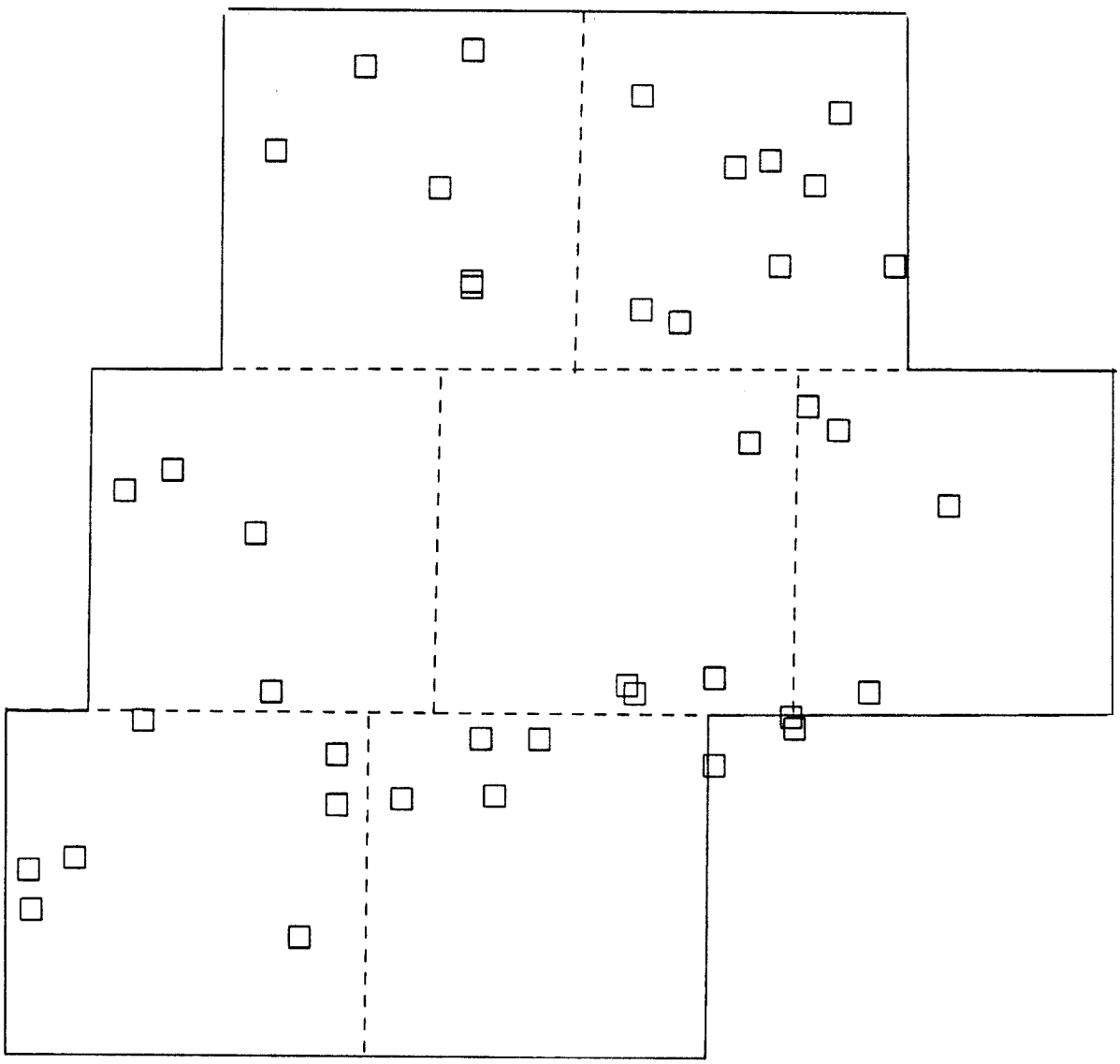


Figure 9.

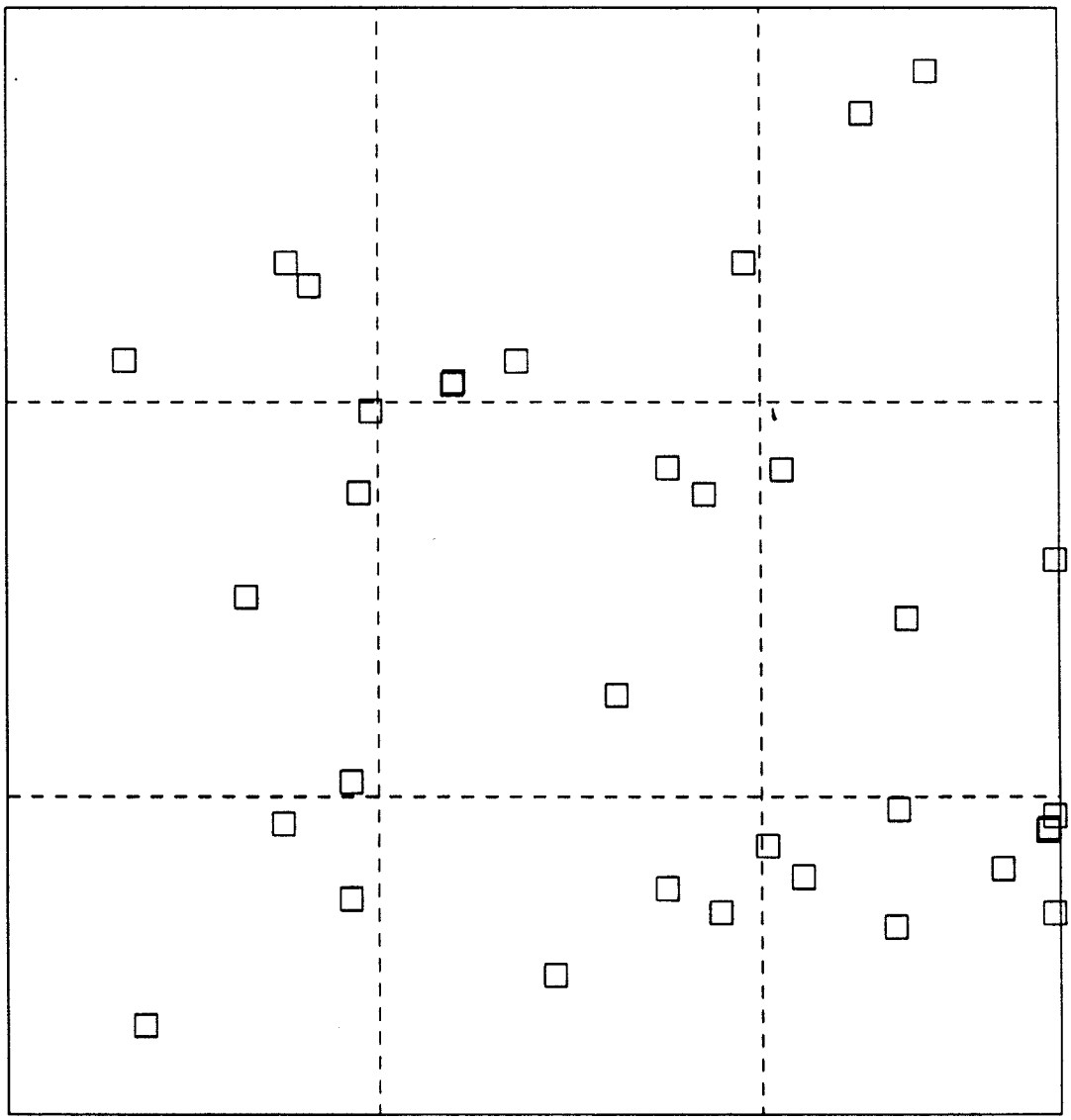


Figure 10.

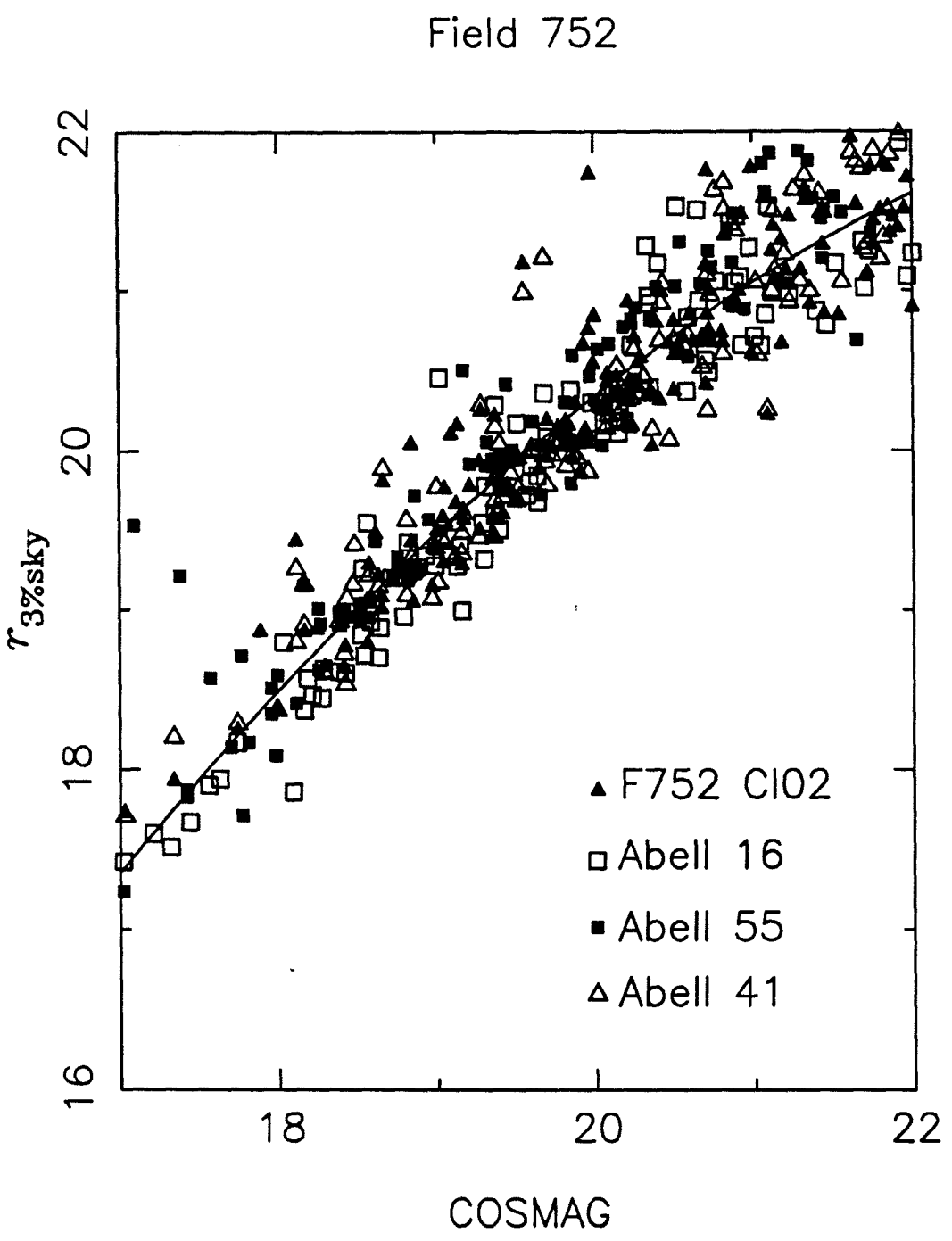

Field 752
22
20
18
16
$r_{3\%sky}$
18
20
22
COSMAG
F752 Cl02
Abell 16
Abell 55
Abell 41

Figure 11.

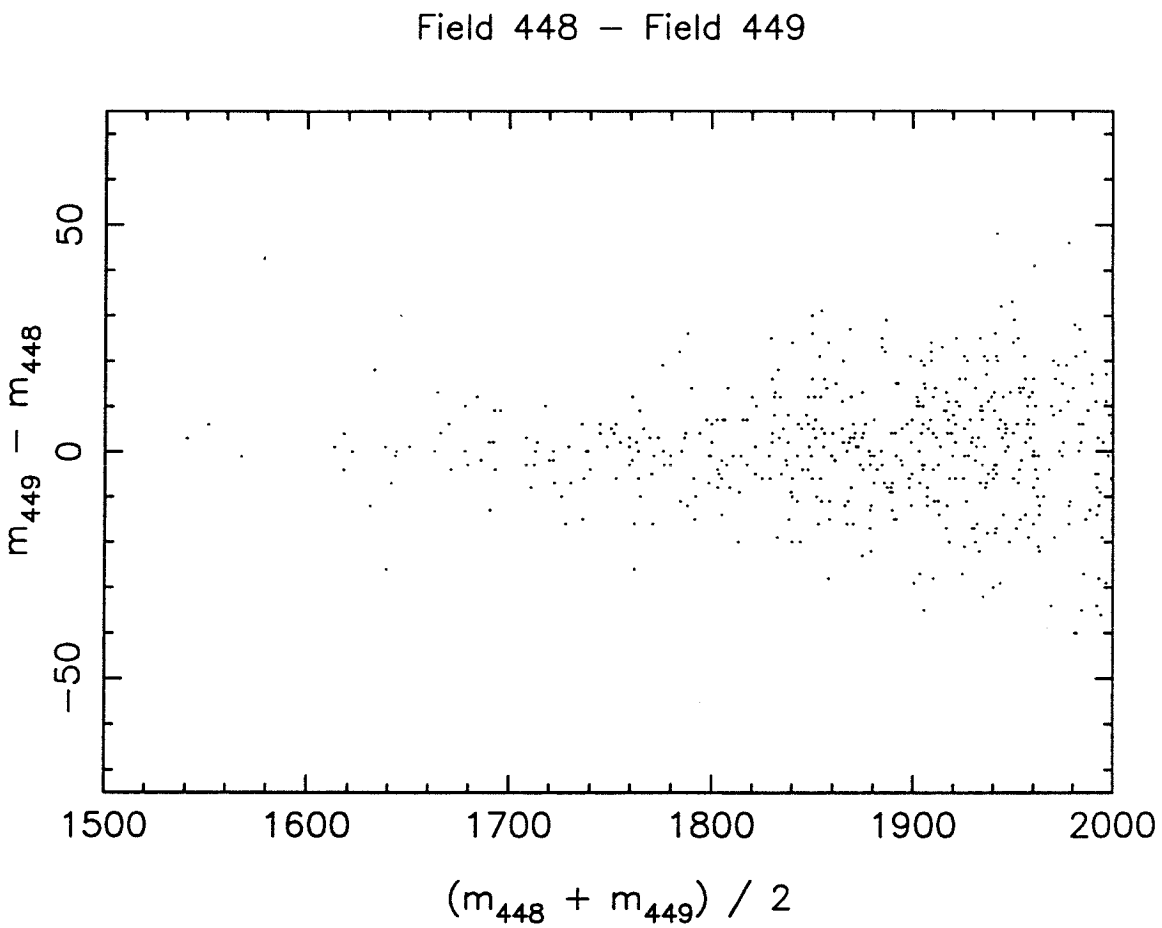

Field 448 – Field 449
50
0
−50
$m_{449} - m_{448}$
1500
1600
1700
1800
1900
2000
$(m_{448} + m_{449}) / 2$

Figure 12.

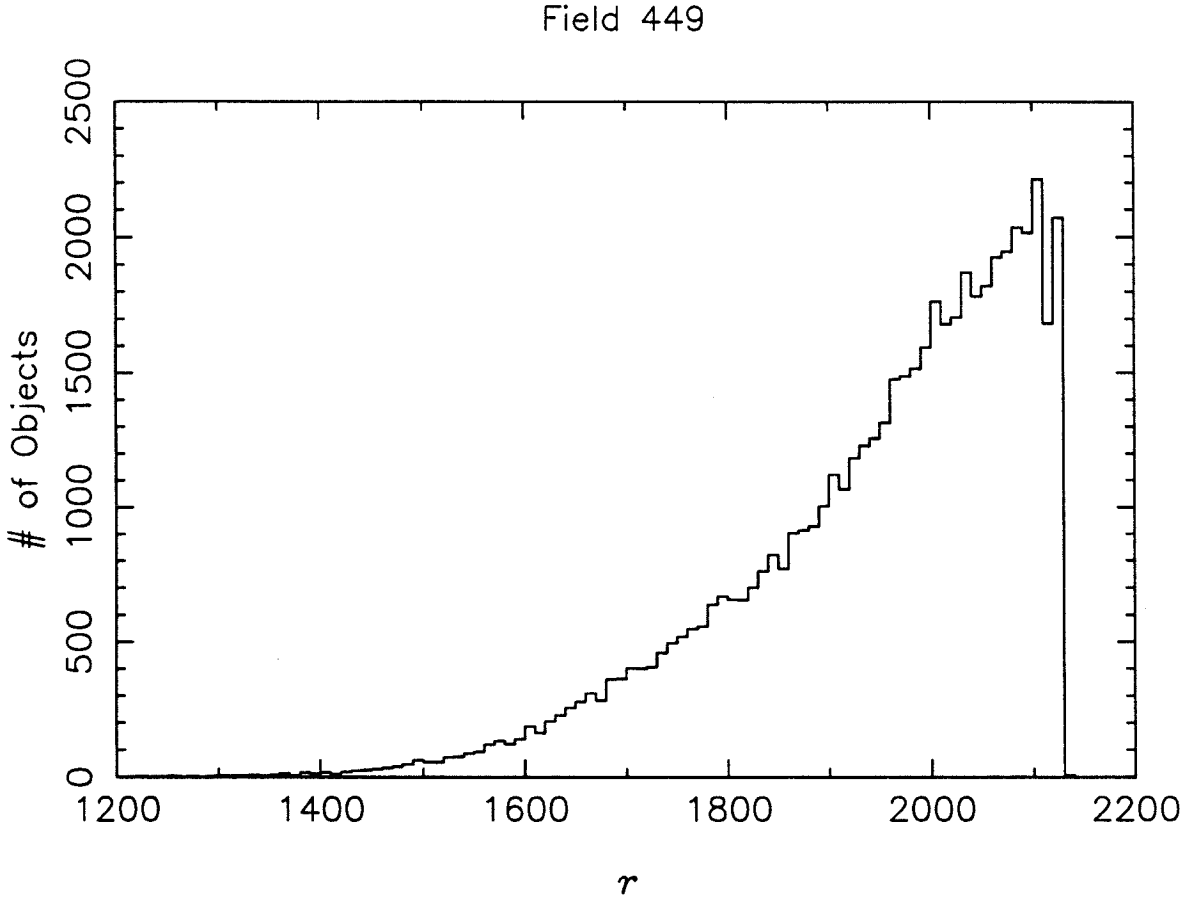

Field 449
# of Objects
0
500
1000
1500
2000
2500
1200
1400
1600
1800
2000
2200
r

Figure 13.

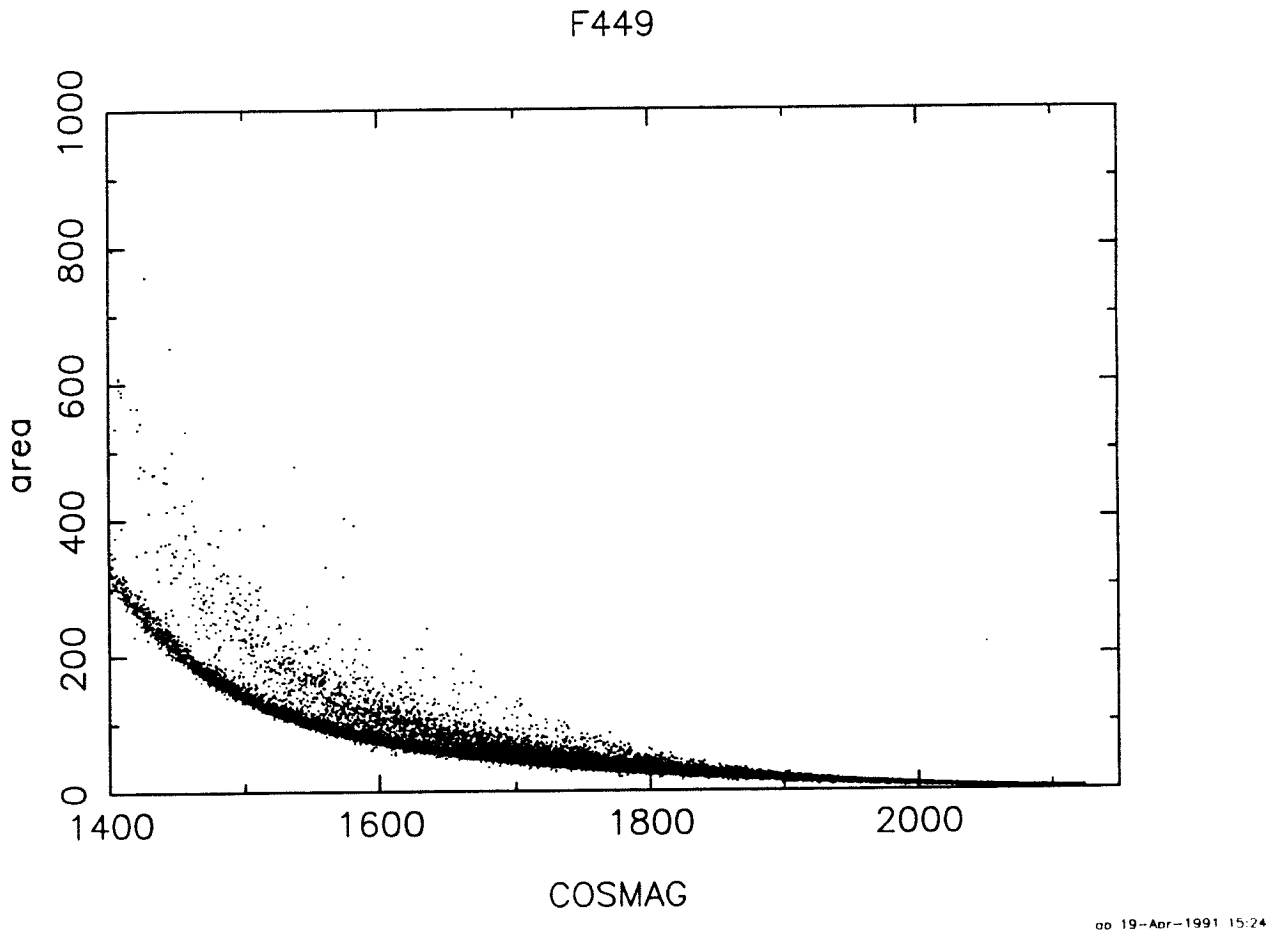

F449
area
0
200
400
600
800
1000
1400
1600
1800
2000
COSMAG
ao 19-Apr-1991 15:24

Figure 15.

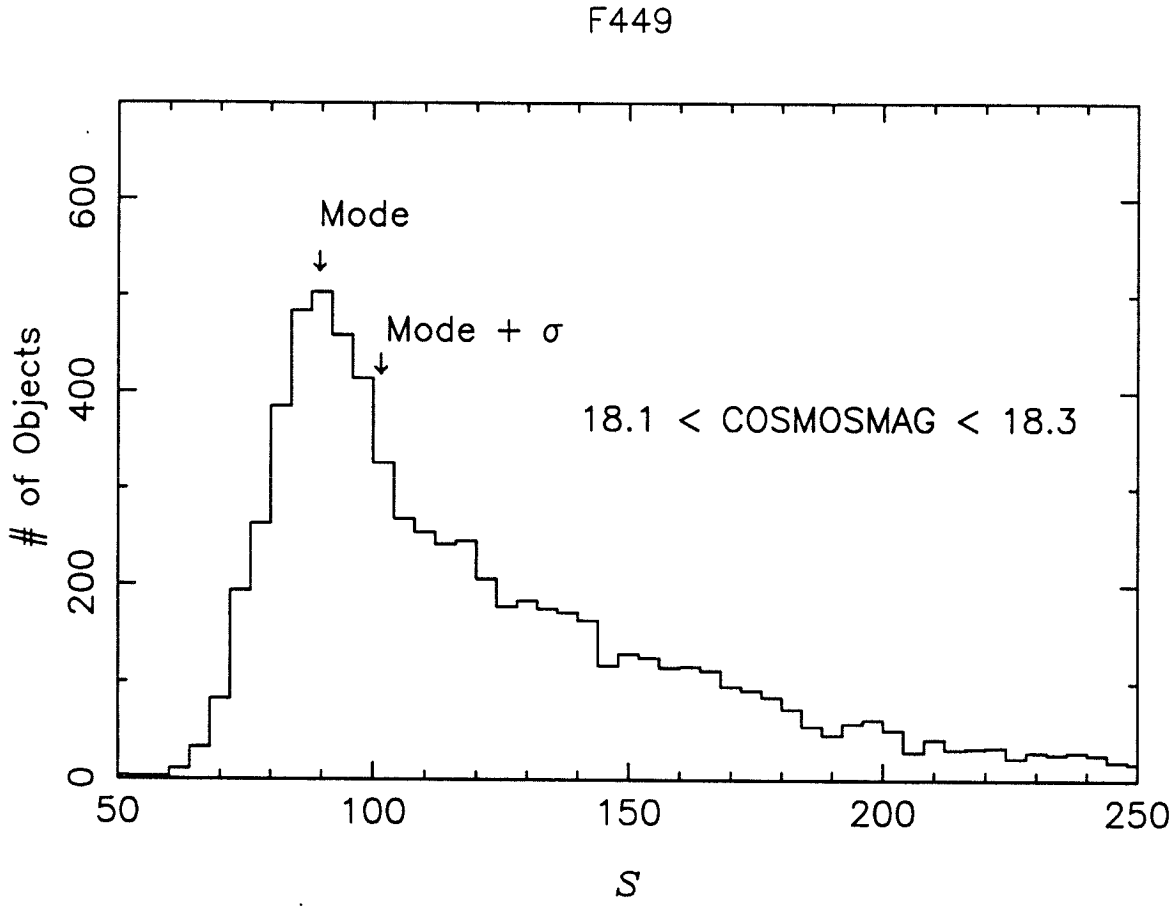

F449
Mode
Mode + σ
18.1 < COSMOSMAG < 18.3
# of Objects
0
200
400
600
50
100
150
200
250
S

Figure 17.

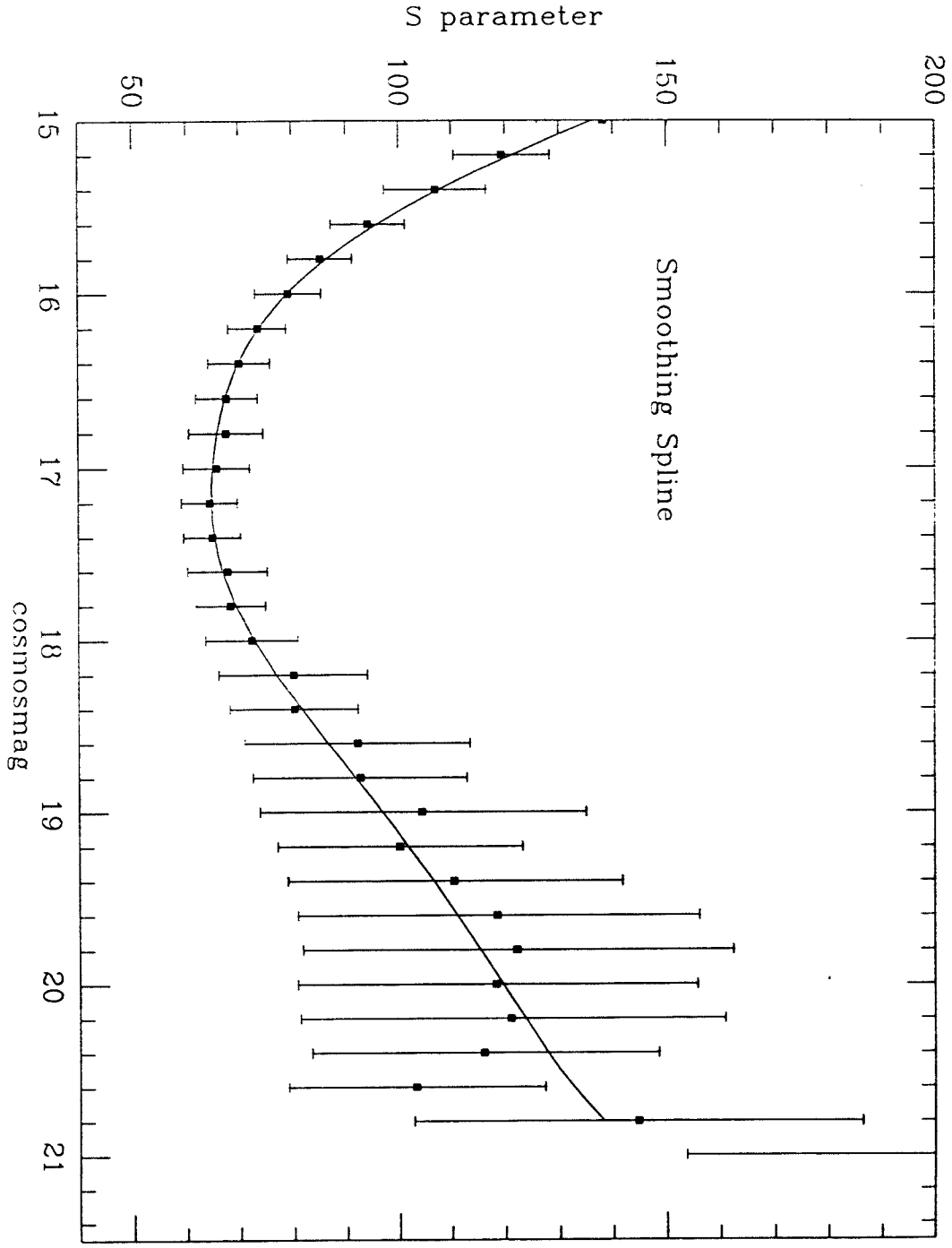

S parameter
50
100
150
200
15
16
17
18
19
20
21
cosmosmag
Smoothing Spline

Figure 16.

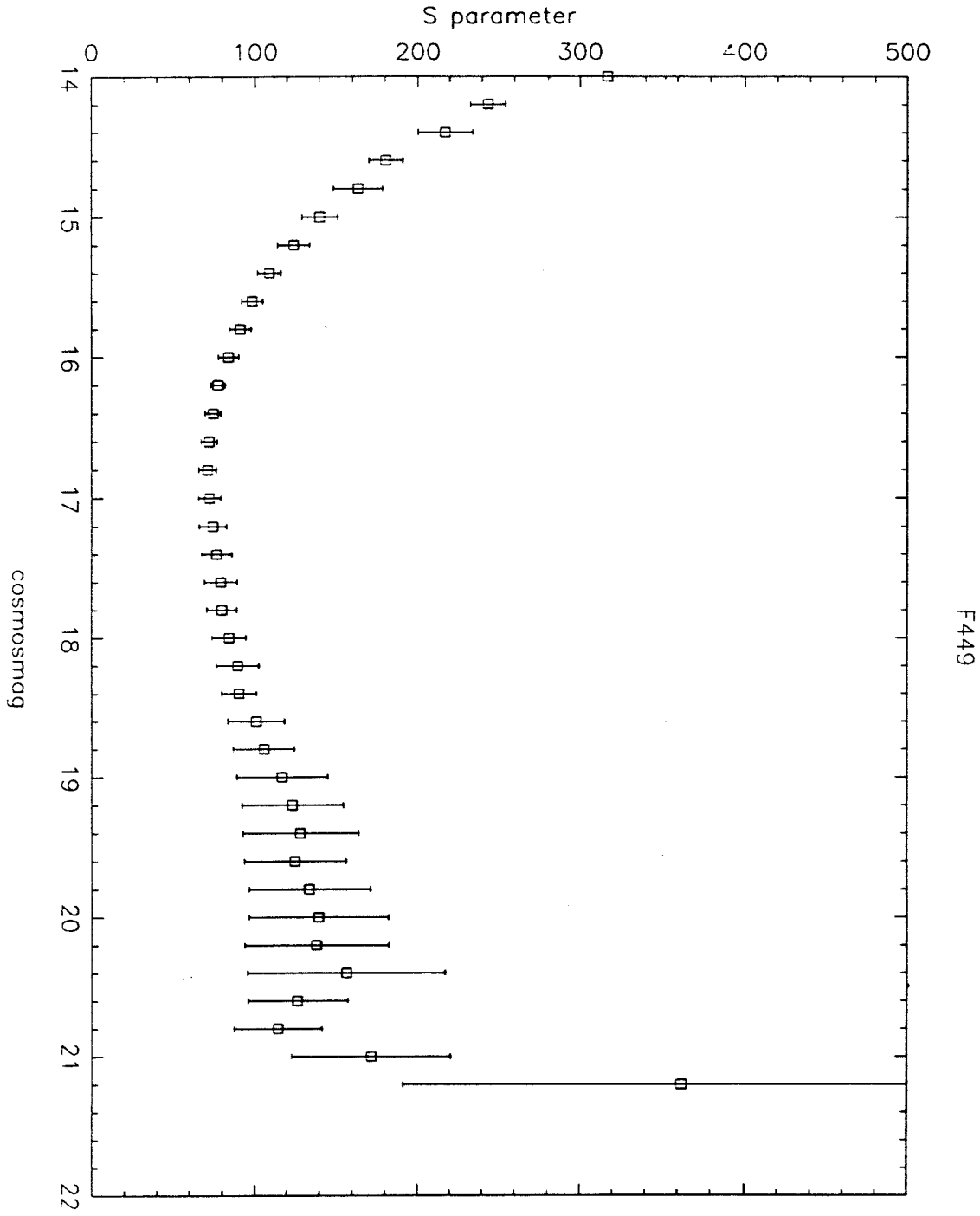

S parameter
0
100
200
300
400
500
14
15
16
17
18
19
20
21
22
cosmosmag
F449

Figure 14.

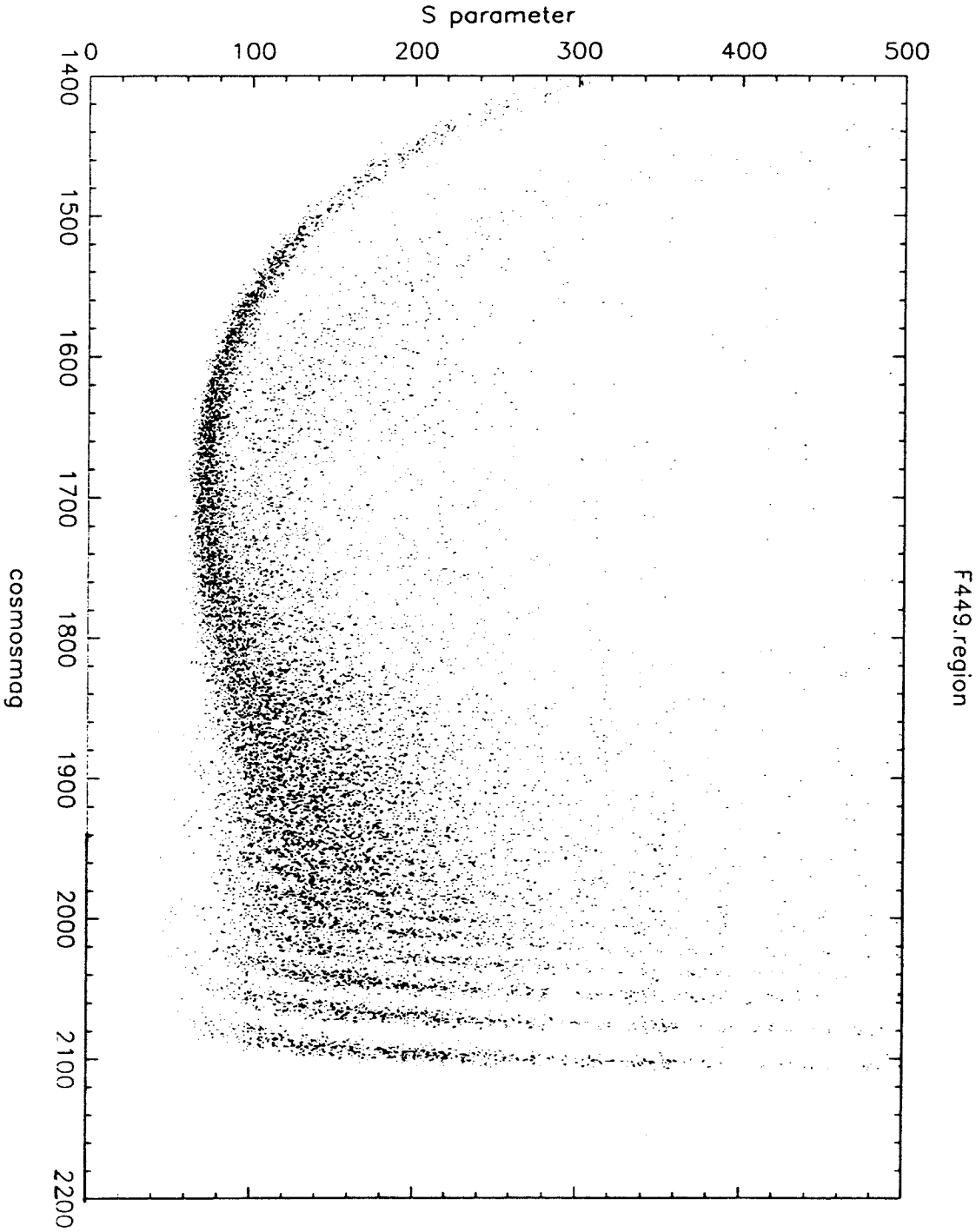

S parameter
0
100
200
300
400
500
1400
1500
1600
1700
1800
1900
2000
2100
2200
cosmosmag
F449.region

Figure 18.

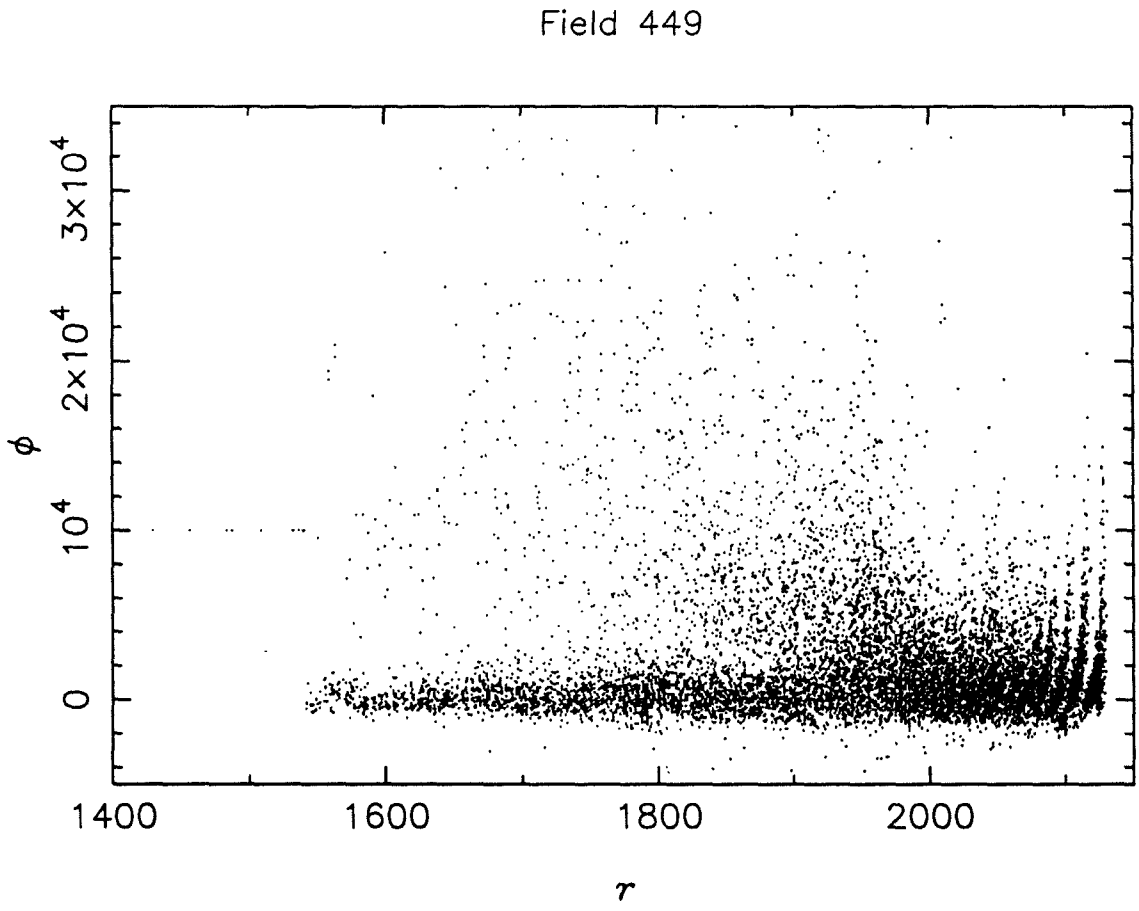

Field 449
ϕ
0
10^4
2×10^4
3×10^4
1400
1600
1800
2000
r

Figure 19.

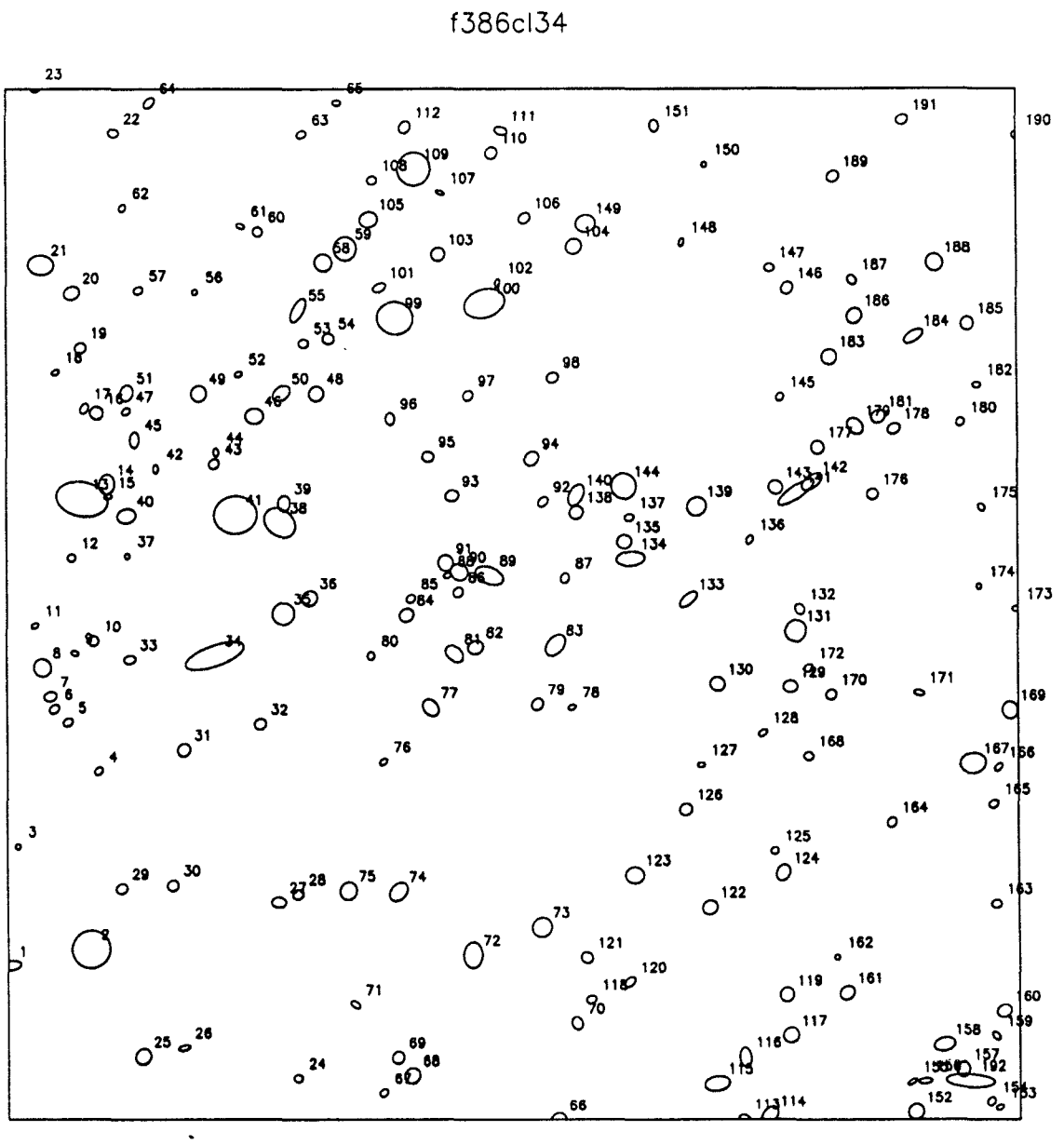


Figure 20.

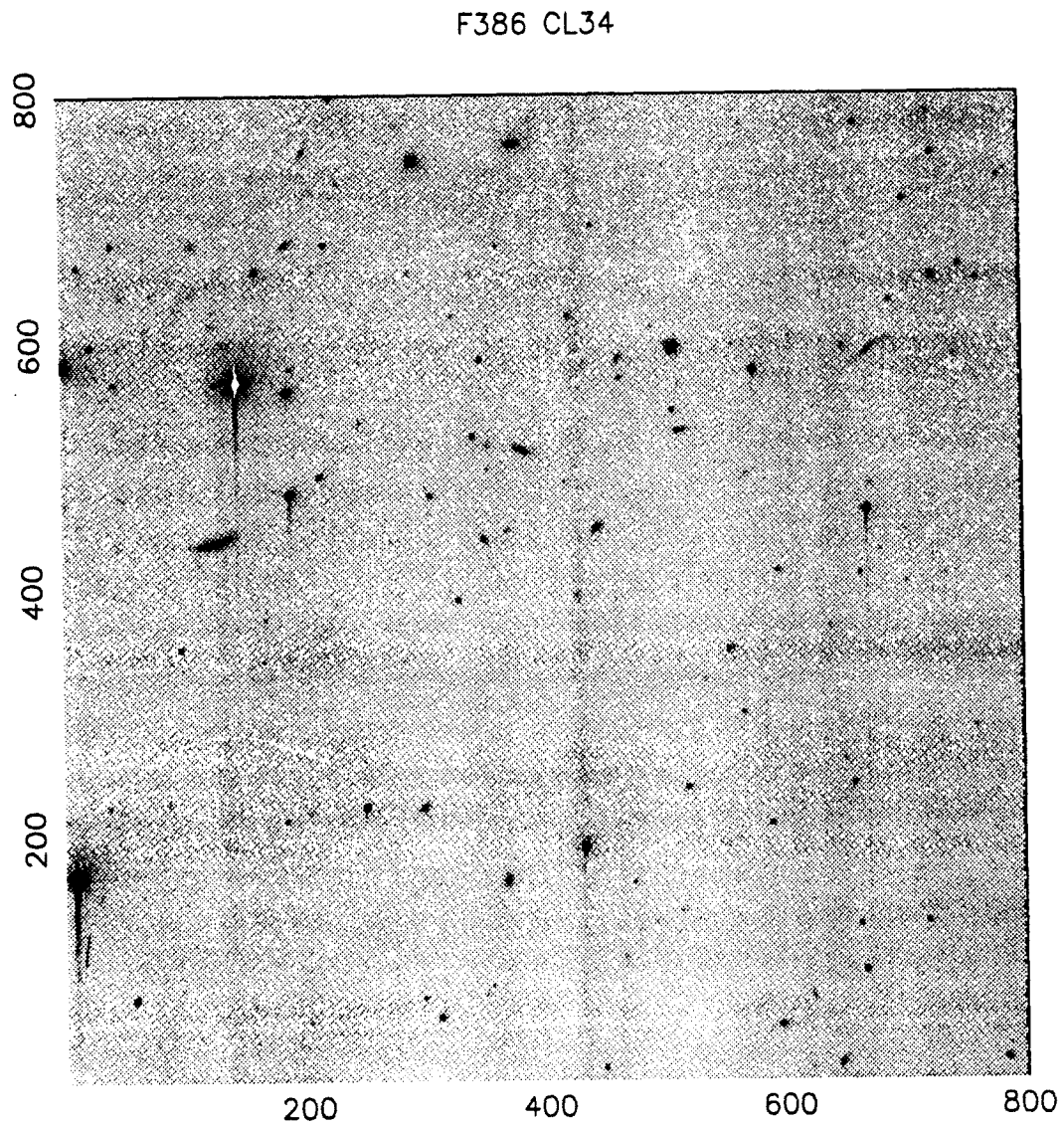


Figure 21.

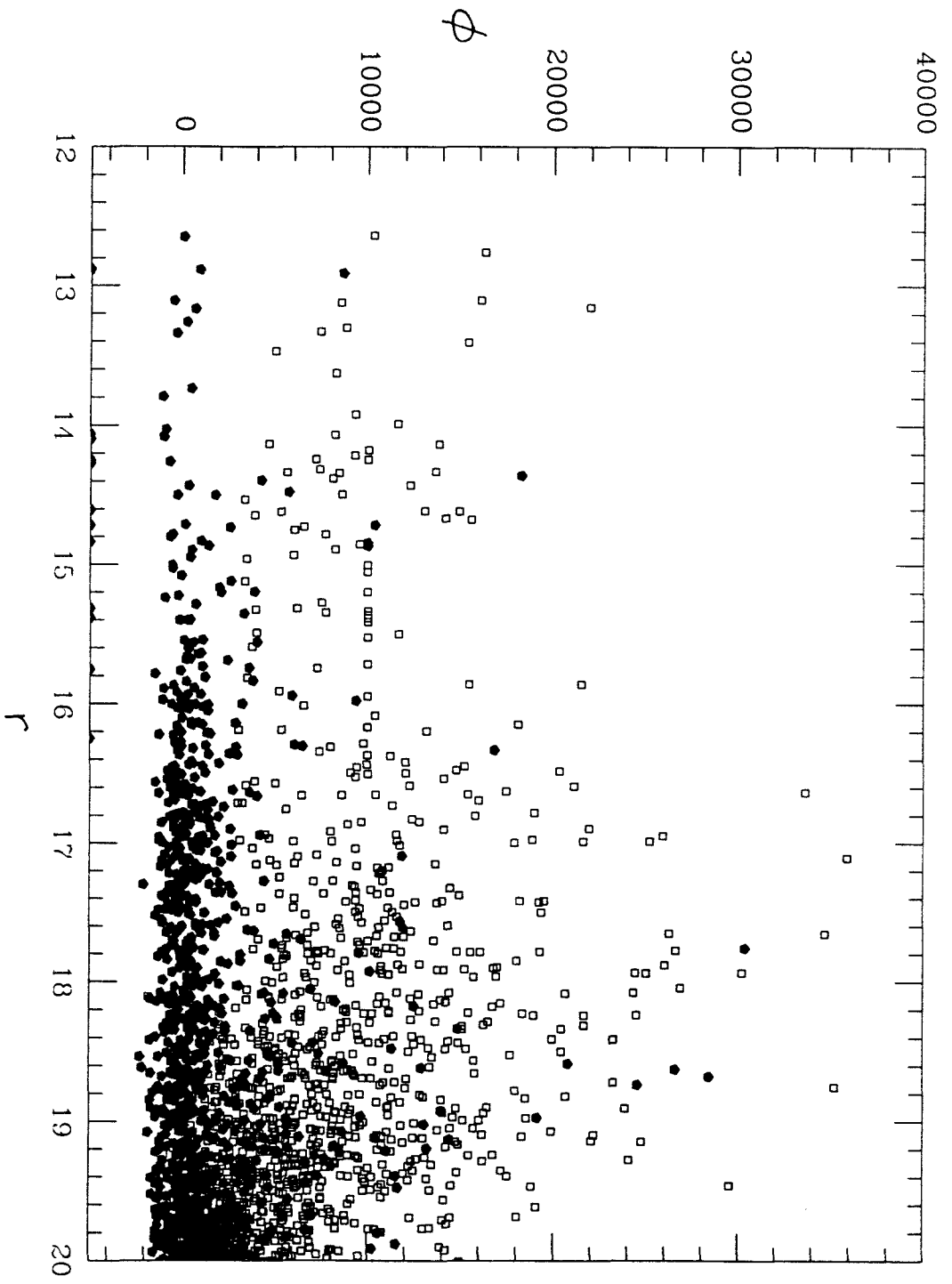


Figure 22.

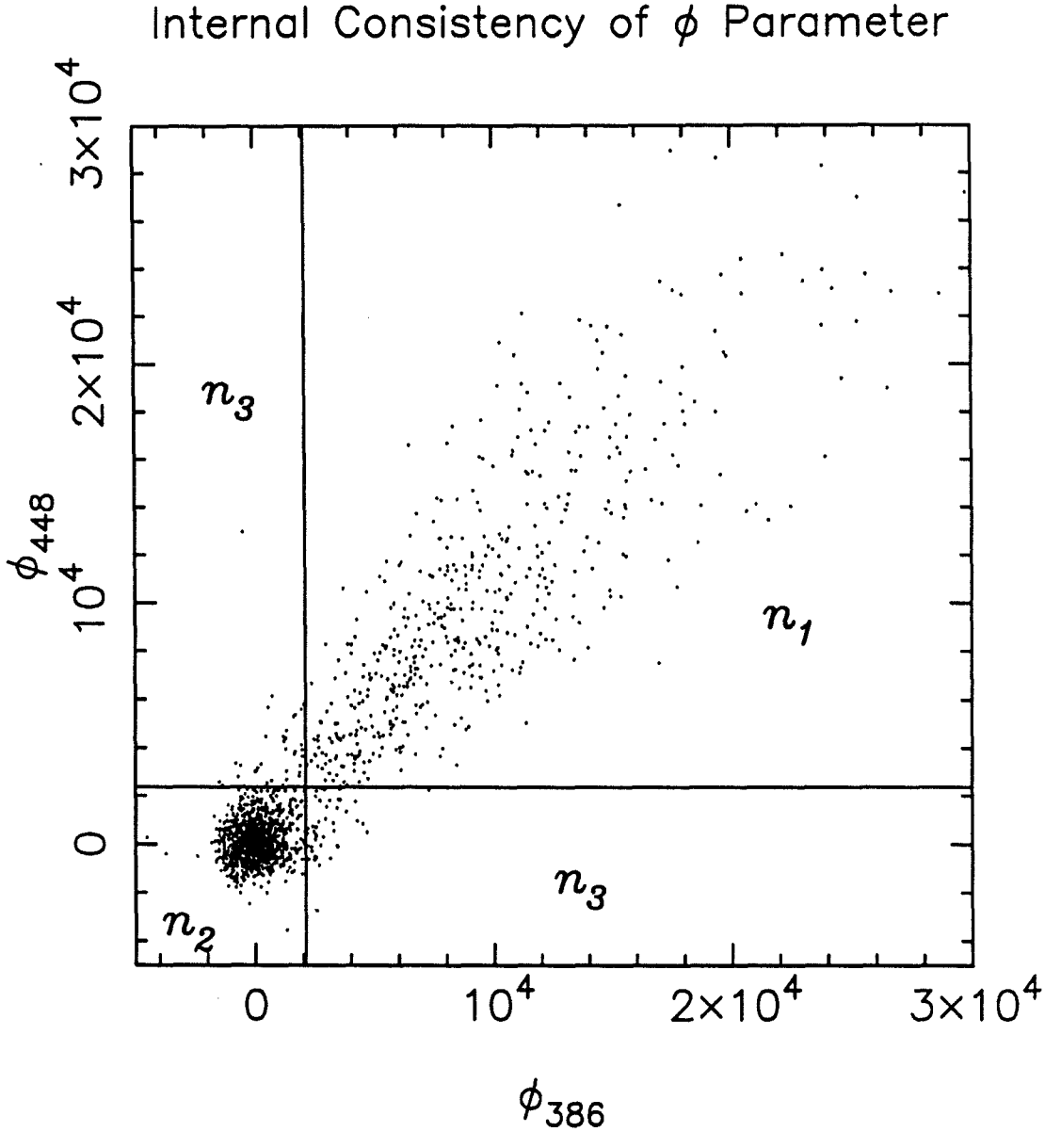

Internal Consistency of ϕ Parameter
ϕ_{448}
ϕ_{386}
0
10^4
2×10^4
3×10^4
n_1
n_2
n_3
n_3

Figure 23.

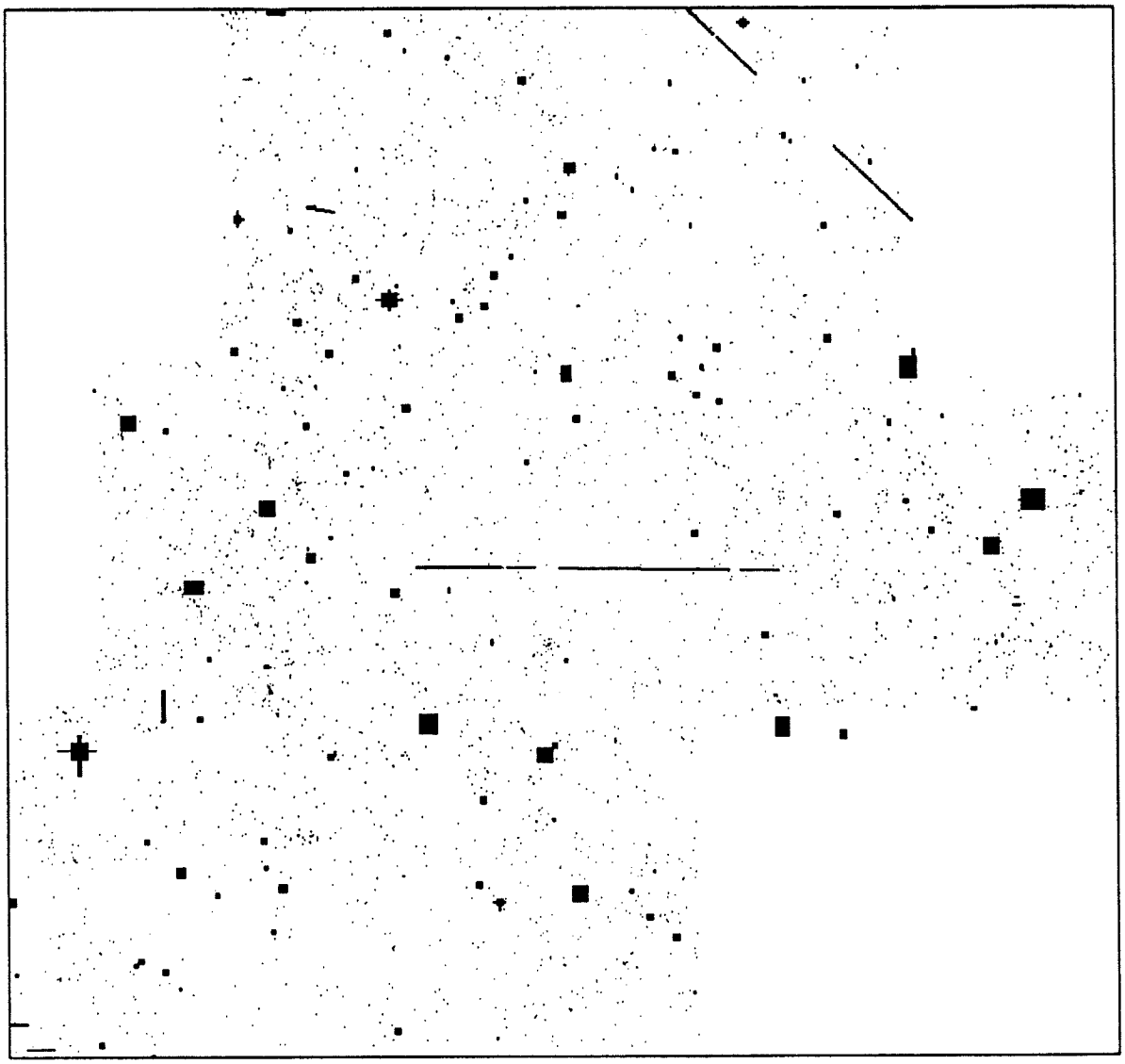


Figure 24.

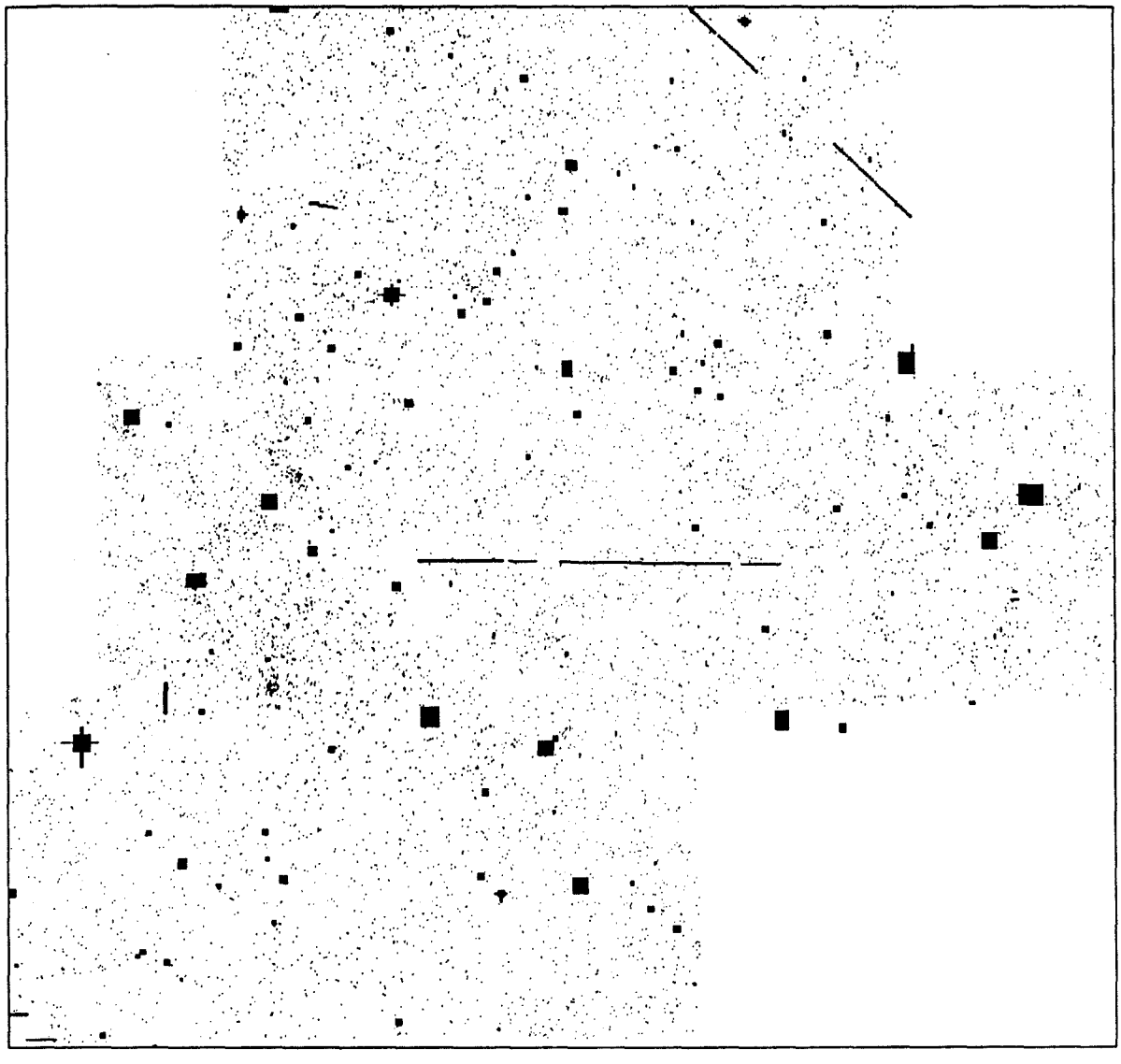


Figure 25.

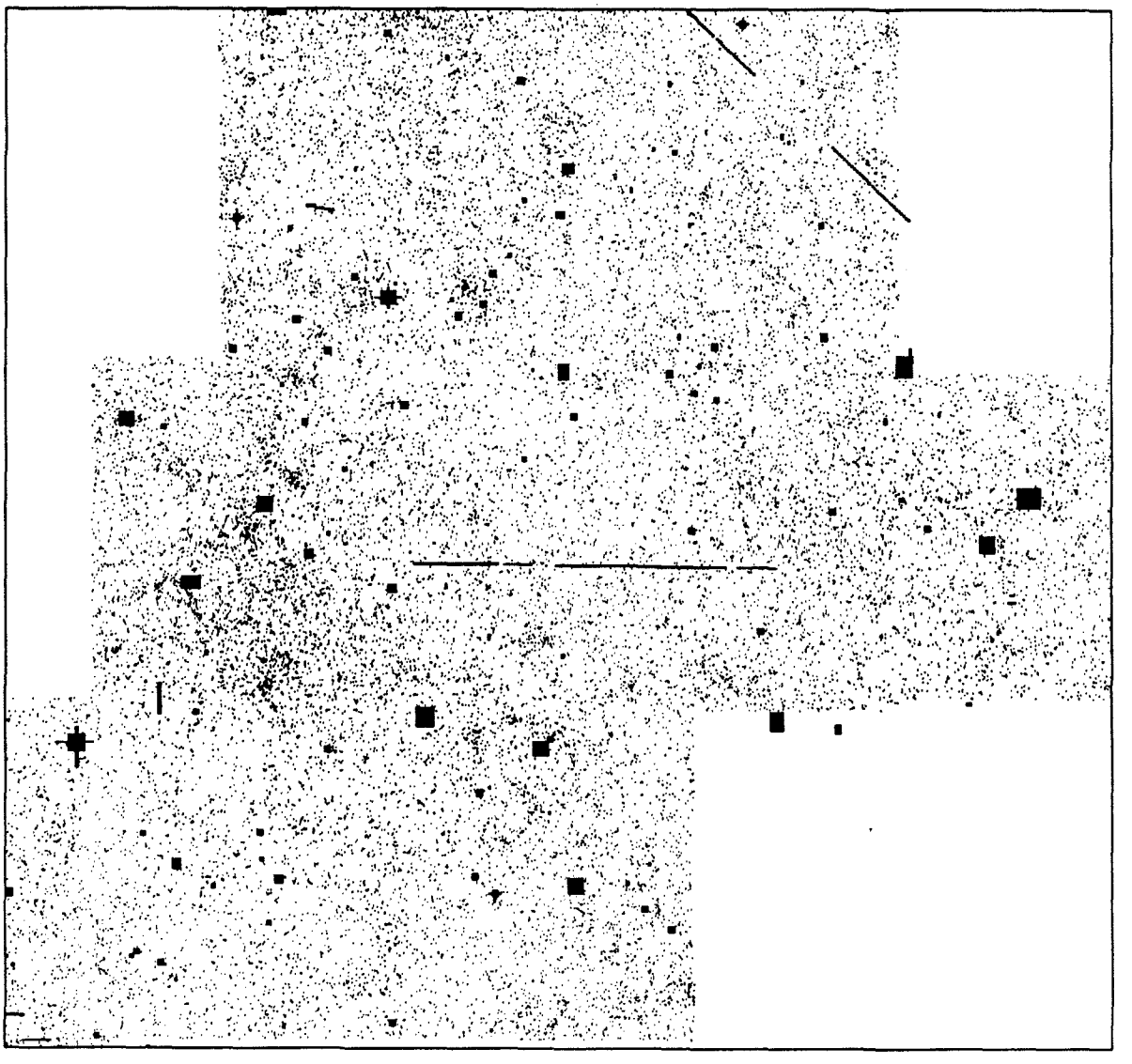


Figure 26.

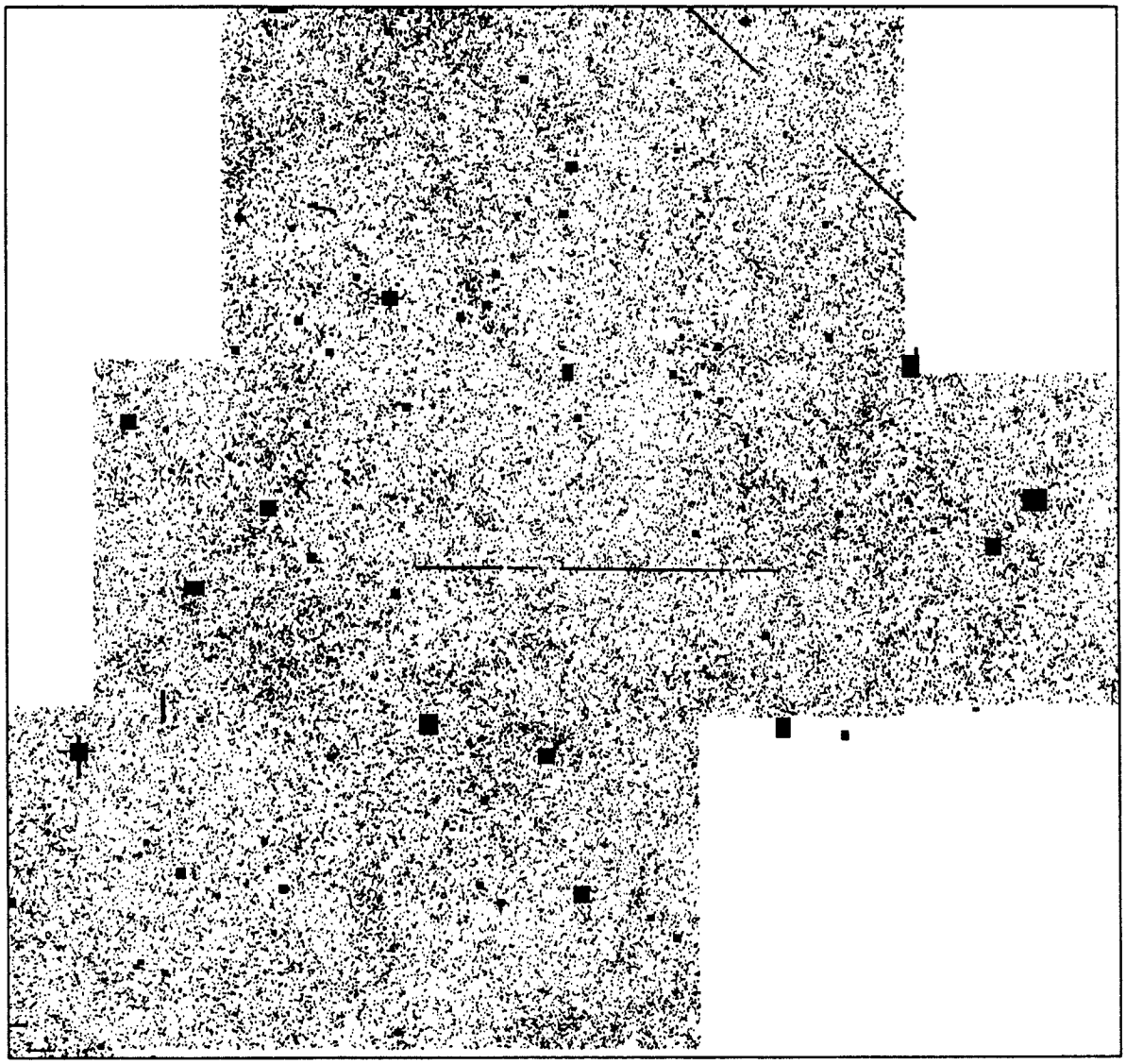


Figure 27.

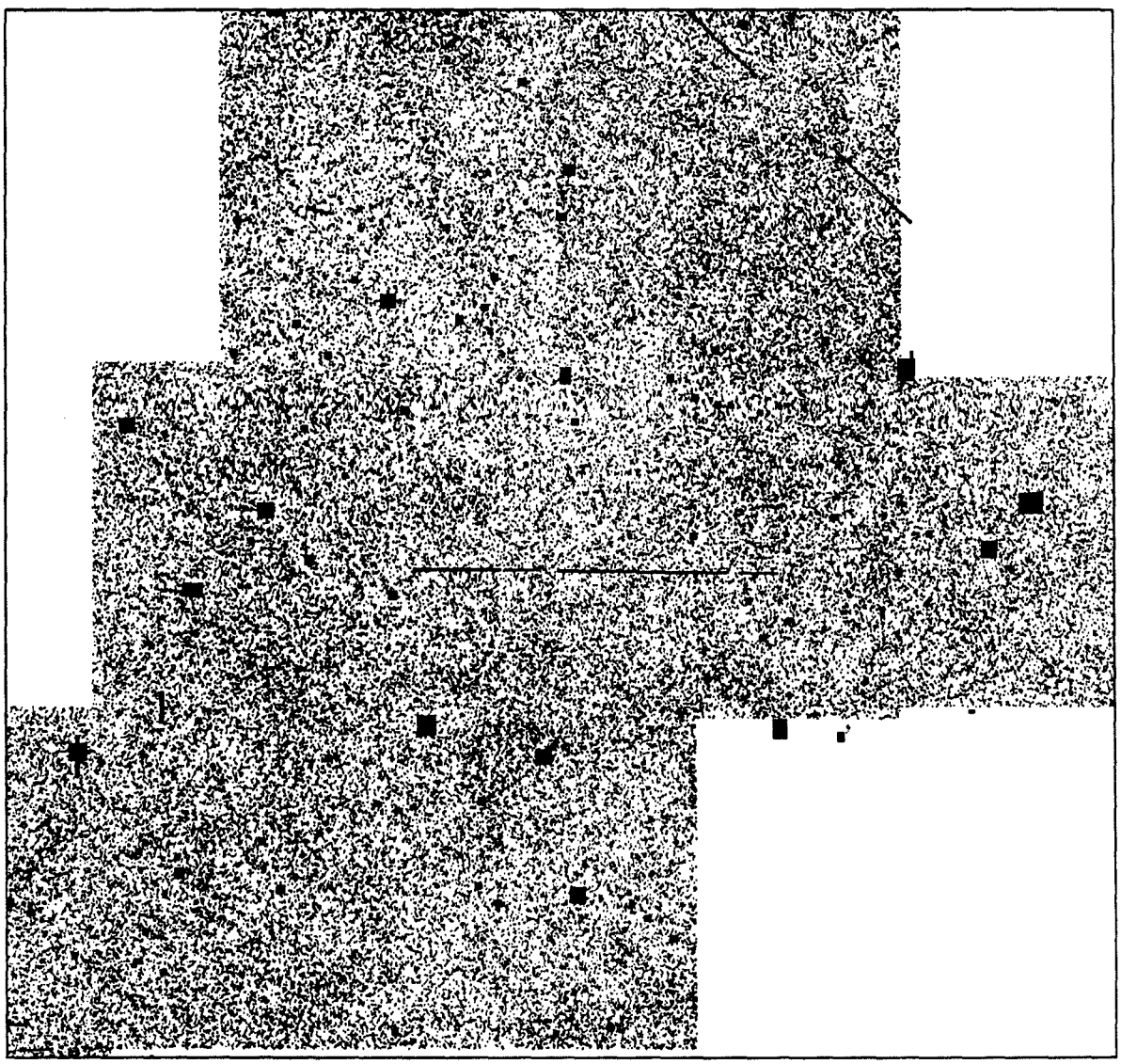


Figure 28.

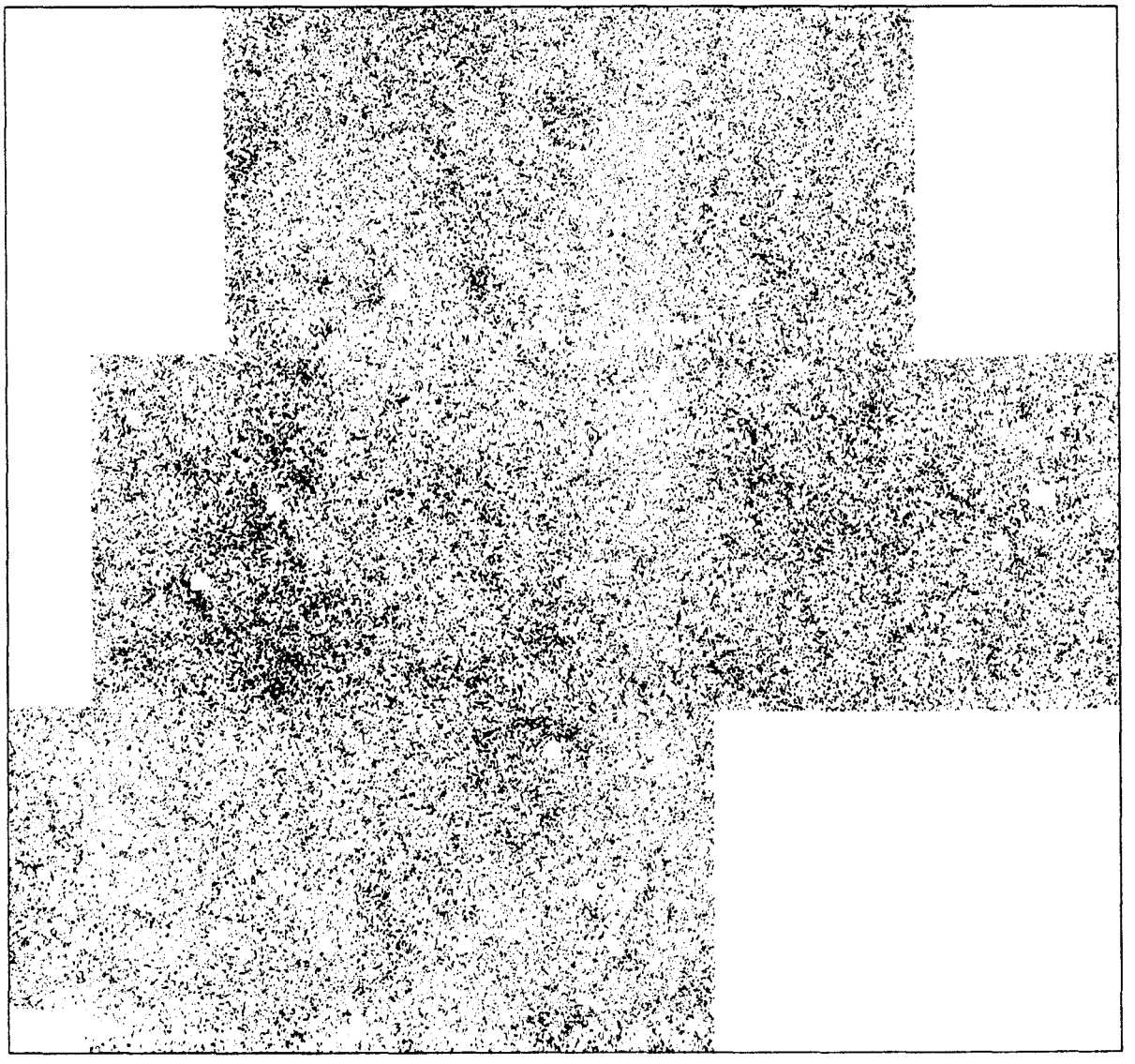


Figure 29.

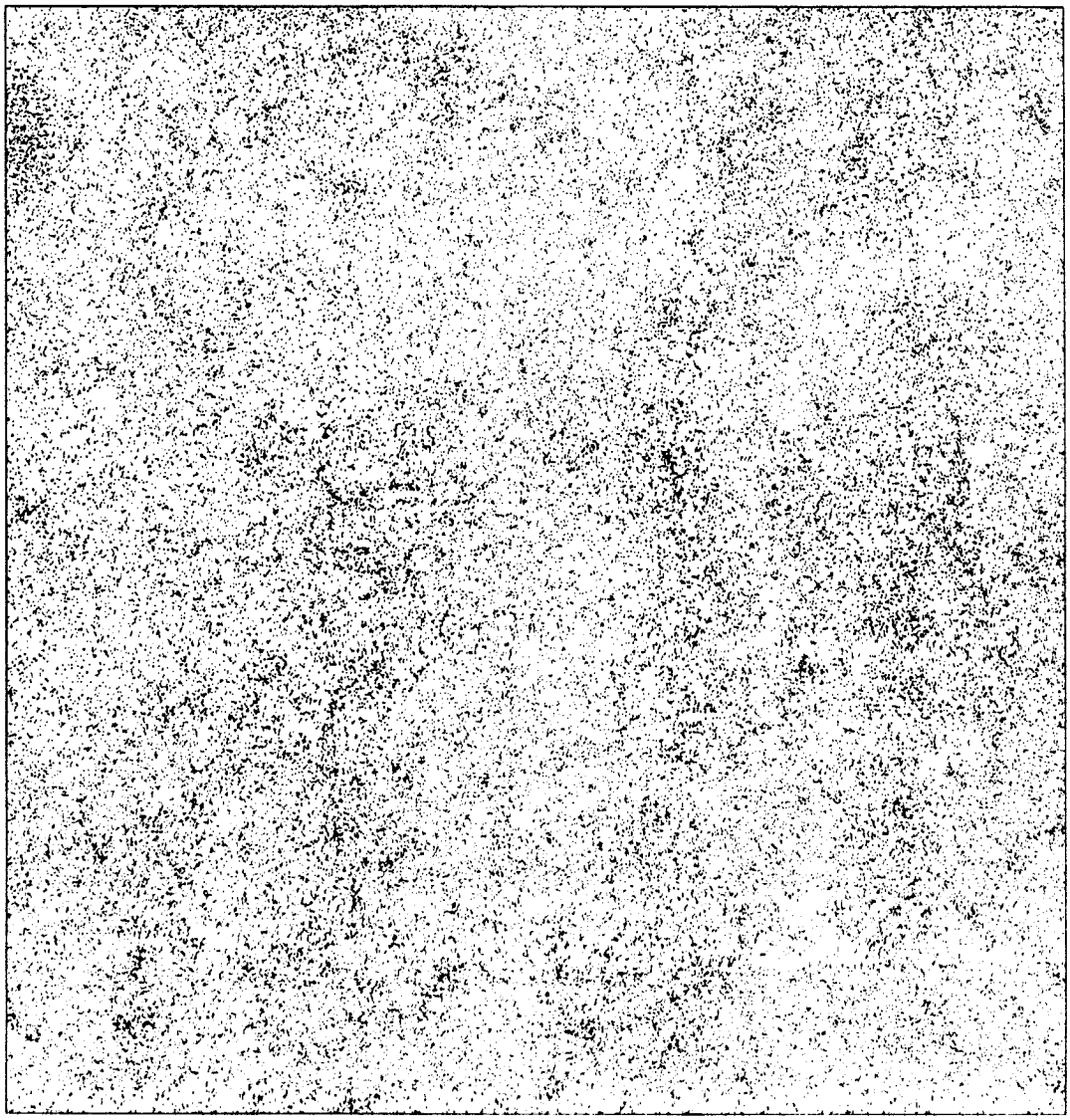


Figure 30.

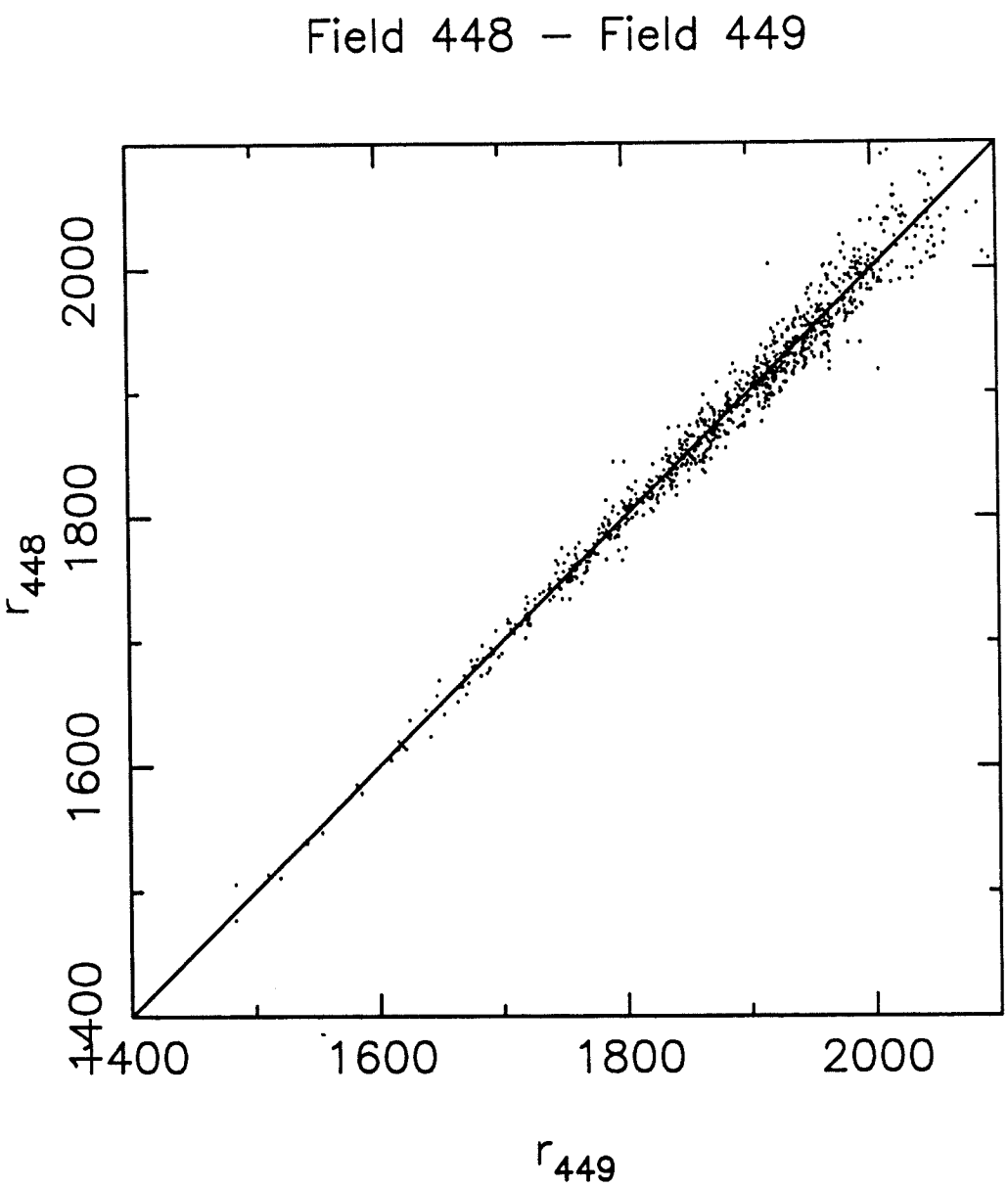

Field 448 – Field 449
r448
r449
1400
1600
1800
2000
1400
1600
1800
2000

Figure 31.

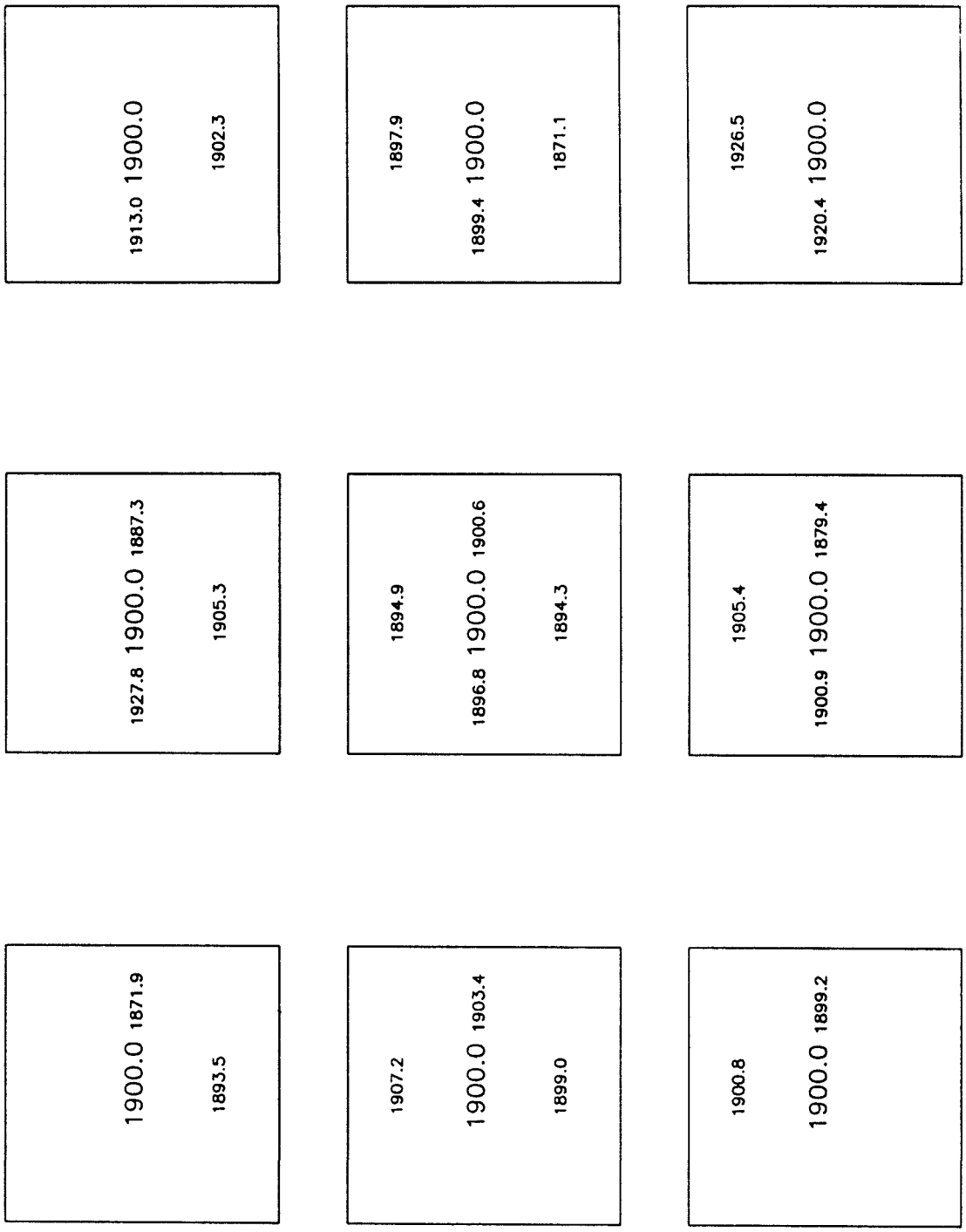

1913.0 1900.0 1902.3
1897.9 1899.4 1900.0 1871.1
1926.5 1920.4 1900.0
1927.8 1900.0 1887.3 1905.3
1894.9 1896.8 1900.0 1900.6 1894.3
1905.4 1900.9 1900.0 1879.4
1900.0 1871.9 1893.5
1907.2 1900.0 1903.4 1899.0
1900.8 1900.0 1899.2

Figure 32.

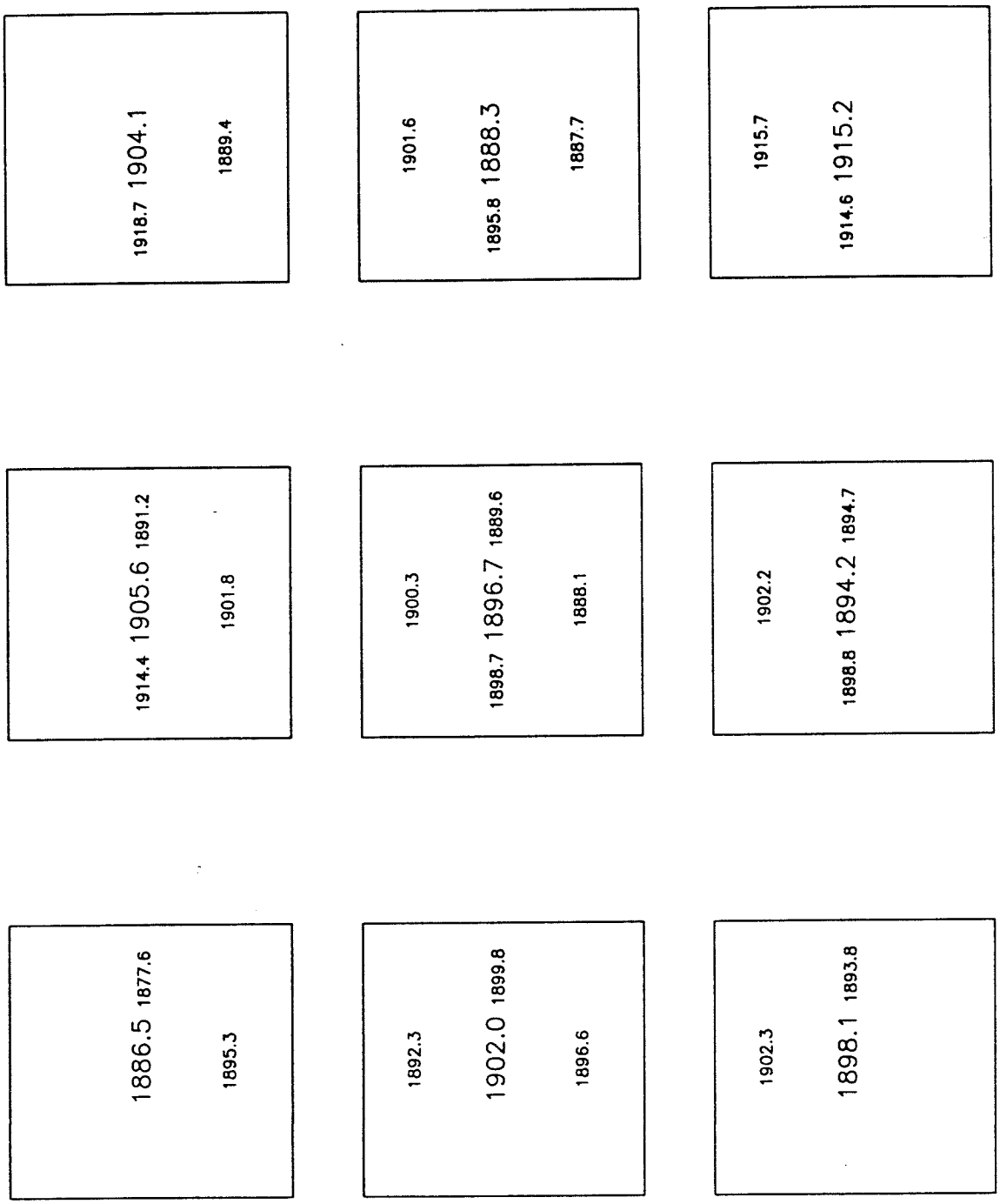

1918.7 1904.1
1889.4
1901.6
1895.8 1888.3
1887.7
1915.7
1914.6 1915.2
1914.4 1905.6 1891.2
1901.8
1900.3
1898.7 1896.7 1889.6
1888.1
1902.2
1898.8 1894.2 1894.7
1886.5 1877.6
1895.3
1892.3
1902.0 1899.8
1896.6
1902.3
1898.1 1893.8

Figure 33.

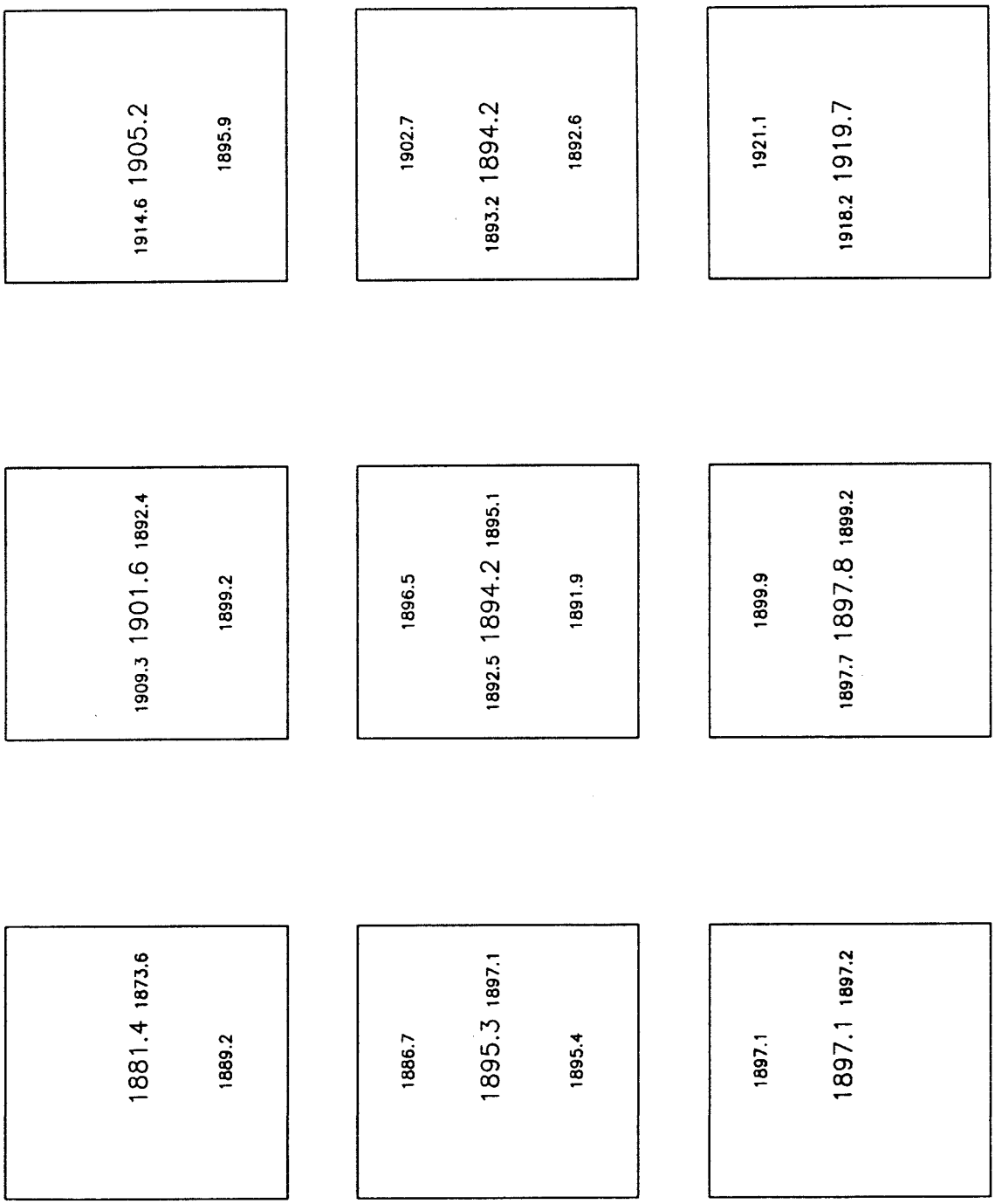

1914.6 1905.2
1895.9
1902.7
1893.2 1894.2
1892.6
1921.1
1918.2 1919.7
1909.3 1901.6 1892.4
1899.2
1896.5
1892.5 1894.2 1895.1
1891.9
1899.9
1897.7 1897.8 1899.2
1881.4 1873.6
1889.2
1886.7
1895.3 1897.1
1895.4
1897.1
1897.1 1897.2

Figure 34.

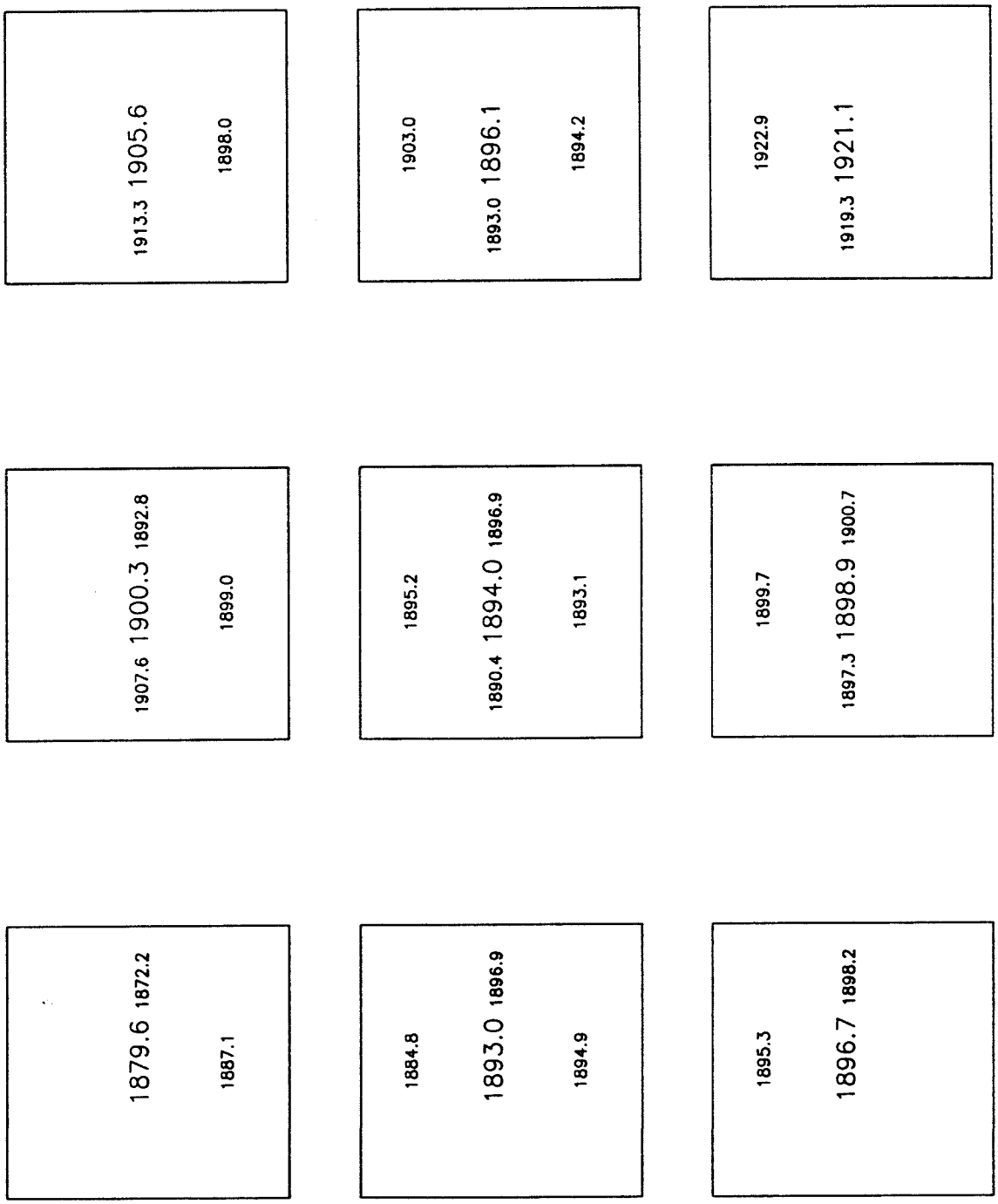
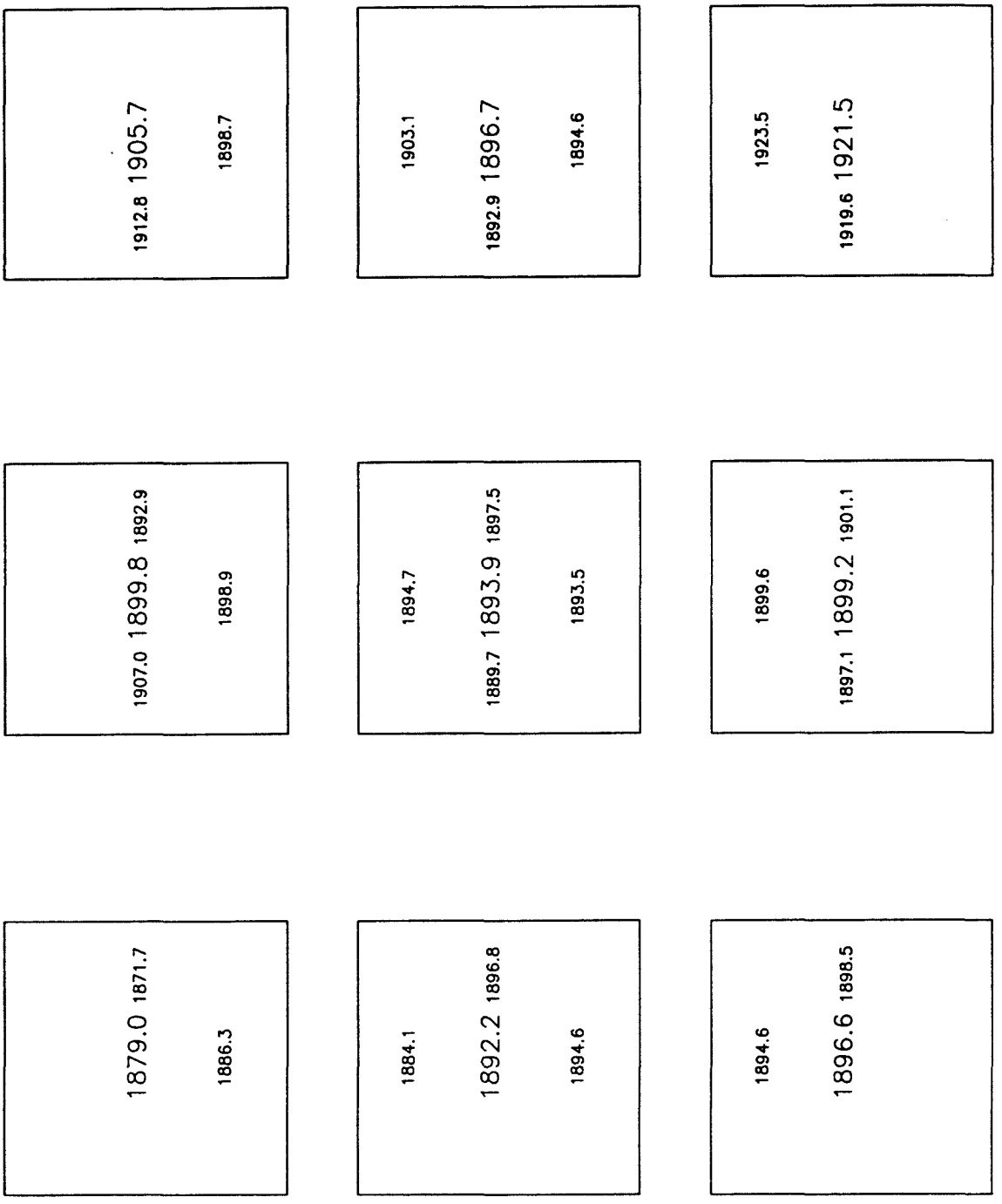

1913.3 1905.6
1898.0
1903.0
1893.0 1896.1
1894.2
1922.9
1919.3 1921.1
1907.6 1900.3 1892.8
1899.0
1895.2
1890.4 1894.0 1896.9
1893.1
1899.7
1897.3 1898.9 1900.7
1879.6 1872.2
1887.1
1884.8
1893.0 1896.9
1894.9
1895.3
1896.7 1898.2

Figure 35.


1912.8 1905.7 1898.7
1903.1 1892.9 1896.7 1894.6
1923.5 1919.6 1921.5
1907.0 1899.8 1892.9 1898.9
1894.7 1889.7 1893.9 1897.5 1893.5
1899.6 1897.1 1899.2 1901.1
1879.0 1871.7 1886.3
1884.1 1892.2 1896.8 1894.6
1894.6 1896.6 1898.5

2.11 References

Abell, G. O. 1958, *Ph.D. thesis, Calif. Inst. of Tech.*

Beard, S. M., MacGillivray, H. T. and Thanisch, P. F., 1990, *M.N.R.A.S.*, **247**, 311.

Collins, C. A., Heydon-Dumbleton, N. H. and MacGillivray, H. T., 1989, *M.N.R.A.S.*, **236**, 7p.

Geller, M. J., de Lapparent, V. and Kurtz, M. J., 1984, *Astrophys. J. (Letters)*, **287**, L55.

Jarvis, J. F. and Tyson, A.J., 1979, *SPIE Proc. Instrumentation in Astronomy*, **3**.

Kent, S. M., 1985, *Pub. Aston. Soc. Pac.*, **97**, 165.

de Lapparent, V., Geller, M. J., and Huchra, J. P., 1988, *Astrophys. J.*, **332**, 44.

MacGillivray, H. T. and Stobie, R. S., 1984, *Vistas in Astronomy*, **27**,433.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989, *M.N.R.A.S.*, **242**, 43p.

Meinhold, P. and Lubin, P., 1991, *in press.*

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling., W. T., 1986, *Numerical Recipes*, Cambridge University Press

Reid, I. N., Brewer, C., Brucato, R. J., McKinley, W. R., Maury, A., Mendenhall, D., Mould, J. R., Mueller, J., Neugebauer, G., Phinney, J., Sargent, W. L. W., Schombert J. and Thicksten, R., 1991, *The Second Palomar Sky Survey* , submitted to *P.A.S.P.*

Schombert, J., 1988, *Astron. J.*, **95**, 1389.

Shortridge 1984, *The FIGARO Users' Guide, unpublished.*

Stetson, P. B., 1987, *Pub. Aston. Soc. Pac.*, **99**, 101.

Thuan, T. X. and Gunn, J. E. G., 1976, *Pub. Aston. Soc. Pac.*, **88**, 543.

Valdes, 1982, *Instrumentation in Astronomy* IV, SPIE **331**,465.

3. The Angular Correlation Function†

3.1 Introduction

With the advent of fast scanning machines, we have witnessed a renewal of interest in statistical analyses of the angular distribution of galaxies, especially the angular two-point correlation function (denoted $w(\theta)$), such as was performed by Groth and Peebles (1977, hereafter GP77) based on a new analysis of the counts of Shane and Wirtanen (1967) (Seldner *et al.* 1977, hereafter SSGP). An objective catalogue of galaxies in the southern hemisphere has recently been completed using the Automated Plate Machine (APM) at Cambridge by Maddox *et al.* (1989a, 1989b), hereafter denoted as MESL, and another has been prepared by Collins *et al.* (1989), using the COSMOS plate scanning machine. Both groups used plates from the United Kingdom Schmidt Telescope (UKST). We have compiled an objective catalogue of galaxies based on plates from the Second Palomar Observatory Sky Survey (POSS-II). It is the first catalogue (in the north) comparable in quality to the survey of MESL. Furthermore, because it is based on independent plate material, and surveys a different region of the sky, it is a powerful test of MESL's results.

Recently, MESL have presented results for the galaxy-galaxy angular two-point correlation function. They find more power on large scales ($\geq 5°$) than previously found by GP77, and argue that SSGP may have removed intrinsic clustering in their analysis. Their result seems at present to be incompatible with standard Cold Dark Matter models with gaussian initial fluctuations.

Collins *et al.* (1989) have also measured $w(\theta)$, but their result disagrees with that obtained by MESL, and is consistent with the measurement of GP77. This is particularly worrisome, as both Collins *et al.* and MESL are using the *same* plate material. The calibration techniques of the two groups are different, as each use their own CCD photometry for the magnitude calibrations, and their own method for doing star-galaxy separation.

† This chapter was published in 1991 in *Astrophysical Journal Letters*, volume 368, page L7.

The APM group attributes the difference between their result and Collins' to the latter's small sample size. Indeed when they create subsets of their catalogue of the same size as that of the COSMOS group and compute $w(\theta)$, they find large variations. These are large enough that the COSMOS result may be just a statistical anomaly. It is crucial to repeat this experiment, preferably using completely independent plate material. This is precisely what we have done in this paper. First, in Section 2, we present an overview of the galaxy catalogue. A complete discussion of the galaxy catalogue will be presented elsewhere† (Picard 1991). In Section 3 we discuss our results for the angular correlation function, which we compare to previous results, and finally present our conclusions.

† This made reference to this work; see Chapter 2.

3.2 The Galaxy Catalogue.

Our data set consists of COSMOS scans of 16 IIIaF plates obtained for the second epoch Palomar Observatory Sky Survey (POSS-II). Each plate covers 6.5° on a side, and the plate centers are offset by 5°. The plates are divided into a north galactic (7 plates, $\alpha \approx 15^h\,0^m, \delta \approx +30°\,(1950.0), b \approx 65°$) and south galactic (9 plates, $\alpha \approx 0^h\,20^m, \delta \approx +5°\,(1950.0), b \approx -55°$) field. We had intended the survey to consist of two regions of 3×3 plates, but two of the plates in the north field had to be omitted because they were of insufficient quality. The COSMOS machine (MacGillivray and Stobie 1984) scans an area of 5.3°on a side, so our overlap regions are 0.3°, and the total area surveyed, after removing bad regions due to bright stars, satellite trails and emulsion defects is 386 square degrees.

The magnitude system of each plate was independently calibrated using CCD photometry in the g and r bands of the photometric system of Thuan and Gunn (1976). This was done because the response of the IIIaF + (RG610 filter) emulsion is very close to that of a Gunn r filter. The CCD data were acquired using a re-imaging camera on the Palomar 1.5 m telescope. Each plate is calibrated with four or more CCD images of clusters of galaxies. All the frames for an individual plate were taken on different nights, so as to avoid possible systematic errors in the magnitude scale due to errors in the zero point variations from night to night. Only the data acquired on photometric nights were accepted, with the requirement that the photometric solution for the night have a scatter $\lesssim 0.035^{\rm m}$. The zero-points in the CCD photometry were determined from observations of standard stars. A typical scatter of these standards about the best fit is $\approx 0.03^{\rm m}$. Standard extinction values were used for the g and r bands (Kent 1985). Our standards confirm these extinction values, so we are confident that there is not a systematic problem in the magnitude systems between the north and south fields, which could arise because the north field is at higher declination, and so the CCD observations are typically taken at lower airmass than those for the south field. Since each plate has at least four CCD fields to calibrate it, this yields a typical error of $0.017^{\rm m}$ for the zero point of each plate's magnitude system due to the error in the zero point of the

photometric calibration. This error includes uncertainties about the atmospheric extinction.

We fit a cubic to the plot of COSMOS magnitudes *vs.* CCD magnitudes, and applied this calibration to all COSMOS magnitudes. The CCD photometry was performed using the Faint Object Classification and Analysis System (FOCAS; Valdes 1982), using an isophote of 24.61 mag per square arc-second. This isophote represents 3% of the typical sky brightness in r at Palomar. Only objects classified as galaxies by FOCAS were used in the magnitude calibrations. Thus the total error in the zero point of a given plate, including the zero point error and the fitting error, is $0.06^{\rm m}$. Such non-uniformities would affect the measurements of $\mathrm{w}(\theta)$ through uncertainties in the mean densities of galaxies. Using a slope for number counts of galaxies of 0.45, one obtains, for the fractional error in surface density of galaxies $\delta N/N = \ln(10)\ 0.45\ \delta(m) = 0.062$, which translates into an uncertainty on $\mathrm{w}(\theta)$ of $(\delta N/N)^2 = 0.004$.

Two tests were performed to ensure the uniformity of the magnitude system across the catalogue. The first was to look for variations in the magnitude scale across an individual plate due to non-uniformities of the emulsion. We verified that such variations were negligible by comparing the magnitude scale derived for each of our CCD frames independently to the final calibration, which was obtained by combining all the available CCD frames for each plate. The second test was to measure the uniformity of the magnitudes across plate boundaries. This was done by comparing the calibrated magnitudes for objects lying in the overlap regions. For a set of plates that are perfectly calibrated and uniform, we expect $r_i = r_j$ for objects that are present on both plates i and j. This test was performed, and the rms difference between the best fit of plate i *vs.* plate j and the $y = x$ line at $r = 19.0$ was $0.10^{\rm m}$ for the south, and $0.07^{\rm m}$ for the north. These are larger than our predicted errors, so some non-uniformity may be present at the $0.05^{\rm m}$ level. Systematic departures from this line may indicate non-uniformities of the COSMOS magnitudes as the objects approach the edges. A full analysis of the possible vignetting effects, by stacking counts of stars and galaxies will be presented

in Picard, 1991. The results of this analysis show that the largest non-uniformity near the edges is of order $0.07^{\rm m}$, and only affects the outermost 40 arc-minutes of the East sides of the plates. The effect is not large enough to account for the difference between the north and south fields, however, as the result for $w(\theta)$ drops by 0.005 at $\theta = 4°$ when we mask out the edges of the plates, and increases by a similar amount for $1° < \theta < 2°$. This does however show how dangerous it is to calibrate a mosaic of plates relying only on overlap regions, where possible systematic errors may be at their worst.

MESL have had to correct for very large field effects (some at the $0.4^{\rm m}$ level) in the UK Schmidt plate material. Because most of the power on 1°—2°scales comes from galaxies on the same plate, it is important to understand the degree of uniformity of the plates. Fortunately, the POSS-II plates do not suffer from such large field effects. We have checked this by performing the following tests. The first is to compare the relation between image parameters such as *magnitude* and *area* for the central region of a plate with a region in the corner, and then look for shifts in the locus which defines stellar objects. This test was applied to several plates, and with the exception of the north east corner of three plates in the south field, no field effects were found. In these small areas, there is an apparent overdensity of bright ($r < 16.5$) galaxies, as bright stars in this region are mistakenly classified as galaxies. Bright objects (with $r < 16.0$) have saturated cores, and tend to get "deblended" by the COSMOS software, often resulting in poor photometry. There are also few calibrated objects at these bright magnitudes, so the photometric transformations are poorly determined. For all these reasons, we we have restricted our sample to objects with $r > 16.5$. The second test is to consider counts of stars, which we assume are uncorrelated. We have constructed stacks of counts of stars for both our north and south catalogues, and find that the deviations in the counts from Poisson noise correspond to a magnitude non-uniformity of $0.03^{\rm m}$. Note that this test also puts an upper limit on the amount of patchy galactic extinction on scales of 8 arc-minutes to 1°. This is an upper limit on the magnitude non-uniformities, as the smooth differential extinction due to the galaxy is not taken into account by this test. Further details are given elsewhere (Picard, 1991)

The procedure for separating stars and galaxies is described in detail in Picard (1991). Briefly, we establish a parameter (denoted ϕ) which is a function of the measured parameters (*area, peak intensity, sky intensity*) and which is nearly independent of the magnitude of the object. This parameter is essentially the distance of an object from the stellar ridge. A cutoff boundary (denoted $\phi_{\rm cut}$) is computed for each plate in a well determined, objective fashion, and objects with $\phi > \phi_{\rm cut}$ are classified as galaxies. Several tests were performed to measure the accuracy of this procedure, by comparing the COSMOS classifications to those of FOCAS on our CCD frames, and by comparing the classifications of the same objects in the overlap regions, as with the magnitudes. These tests establish that we find 90% of the galaxies at $r = 19.0$. The completeness drops fairly rapidly for fainter objects, so we have restricted our final catalogue to objects with magnitudes between 16.5 and 19.0 .

The depth (D) of our survey relative to that of MESL was calculated by considering the surface density of galaxies (σ) in each survey; from this we obtain $\frac{\rm D_{POSS}}{\rm D_{MESL}} = \left(\frac{\sigma_{\rm POSS}}{\sigma_{\rm MESL}}\right)^{(\frac{1}{3})} = 1.15$.† Our survey is thus 15% deeper than that of MESL.

† For our value of σ, we used an area weighted average of our two fields. The value of σ used for MESL is obtained from their published value at $B_J = 20.0$ and assuming a slope for the galaxy counts of 0.42 .

3.3 The Angular Correlation Function

The angular correlation function w(θ) was computed using the standard estimators described in MESL. For small angles, we use

$$\mathrm{w}(\theta) = F\frac{N_{dd}}{N_{dr}} - 1 \qquad (\theta \leq 0.7^\circ) \tag{1}$$

where N_{dd}, N_{dr} are the numbers of distinct pairs in the data-data and the data-random catalogues respectively, separated by angle $\theta \pm \delta\theta$, and F is the factor which normalizes the number of pairs between the data and the random catalogue to the number of pairs in the data catalogue with itself. The random catalogue is generated with the same geometry as the data catalogue, and contains 100,000 points.

For angles $\theta > 0.7^\circ$, we use an estimator based on counts in cells, as in MESL, to wit,

$$\mathrm{w}(\theta) = \frac{\langle N_i \rangle \langle N_j \rangle}{\langle N_{ij} \rangle} - 1 \qquad (\theta > 0.7^\circ) \tag{2}$$

where N_i, N_j are counts in cells i, j and $\langle\rangle$ denotes the average over all pairs of cells separated by angle $\theta \pm \delta\theta$.

Our results are presented in Figure 1. The filled squares are for the south field, the filled triangles for the north. The results of MESL are also shown for comparison as the open triangles, as are those of Collins *et al.* , shown as crosses. To first approximation, we should scale our correlation function to that of MESL by shifting our measurements to the right and upwards (Peebles 1980) by $\log(\frac{D_{POSS}}{D_{MESL}}) = .06$, which an imperceptible amount on this figure, so we have not done any scaling and simply plotted their results and ours on the same scale. The points of Collins *et al.* are clearly higher than ours at small angle, and indicate that their survey is not as deep as ours. However, we are more interested in the discrepancy at large angles, which is immediately apparent even without scaling, and so none has been performed.

Power laws were fitted to the data for angles less than 0.7°: the slopes obtained are -0.67 ± 0.02 and -0.72 ± 0.01, with amplitudes at $\theta = 1°$ of 0.033 ± 0.001 and 0.0261 ± 0.0004 for the south and north fields respectively. We performed a Monte-Carlo simulation in which we perturbed by 3σ the zero point of the magnitude scale and the choice of $\phi_{\rm cut}$ for each plate. The error in magnitude introduced is about the size of the tolerance which Geller *et al.* (1984) showed was necessary to properly measure $w(\theta)$ on scales larger than one plate. We use the difference between the correlation function obtained with this perturbed catalogue and the original catalogue to estimate the size of the error bars due to systematic errors in the calibrations of the plates. These are the error bars plotted in Figure 1.

3.4 Conclusions

The most striking feature on Figure 1 is the significant disagreement between the north and south fields on scales larger than 0.5°. This disagreement is larger than can be attributed to systematic errors in our calibration. We must therefore conclude that this discrepancy between the two fields is an indication of the difference in the large scale structure between regions of this size.

We find evidence for clustering on large scales as strong as indicated by MESL. Even in our north field, the correlation function does not drop significantly below that of MESL for angles less than 2°, which corresponds to approximately 4° at the depth of the Lick survey. Thus our results confirm the result of MESL that there is more power on large scales than seen by GP77. The high accuracy of our calibrations eliminates the need for filtering our galaxy map, which is dangerous because this process can remove power on large angles.

This strong clustering on large scales can be seen even on a sample of modest size. Our sample is comparable in size to the "small" sample of Collins *et al.* The discrepancy between our respective results† (they obtain a correlation function which is consistent with that of GP77) is interesting. Given the large difference between our north and south fields, the discrepancy might be the result of a (real) difference in the power of the fields examined. But it is possible that systematic calibration errors are to blame, as it is difficult to eliminate the position dependent errors from field effects in their magnitude scale and star-galaxy separation. Also, as discussed in Section 2, it is dangerous to calibrate the plate to plate magnitude scale using the overlap regions, which is what Collins *et al.* have had to do, having only 3 CCD frames for 6 plates. We have not relied on overlap regions for calibrations, as we have four CCD frames per plate.

Obviously, the large and significant differences in the degree of galaxy clustering between our north and south fields, obtained with the most direct and extensive

† Since the publication of this paper, the COSMOS group has enlarged its survey and now find good agreement with our results (Nichols, *priv. comm.*).

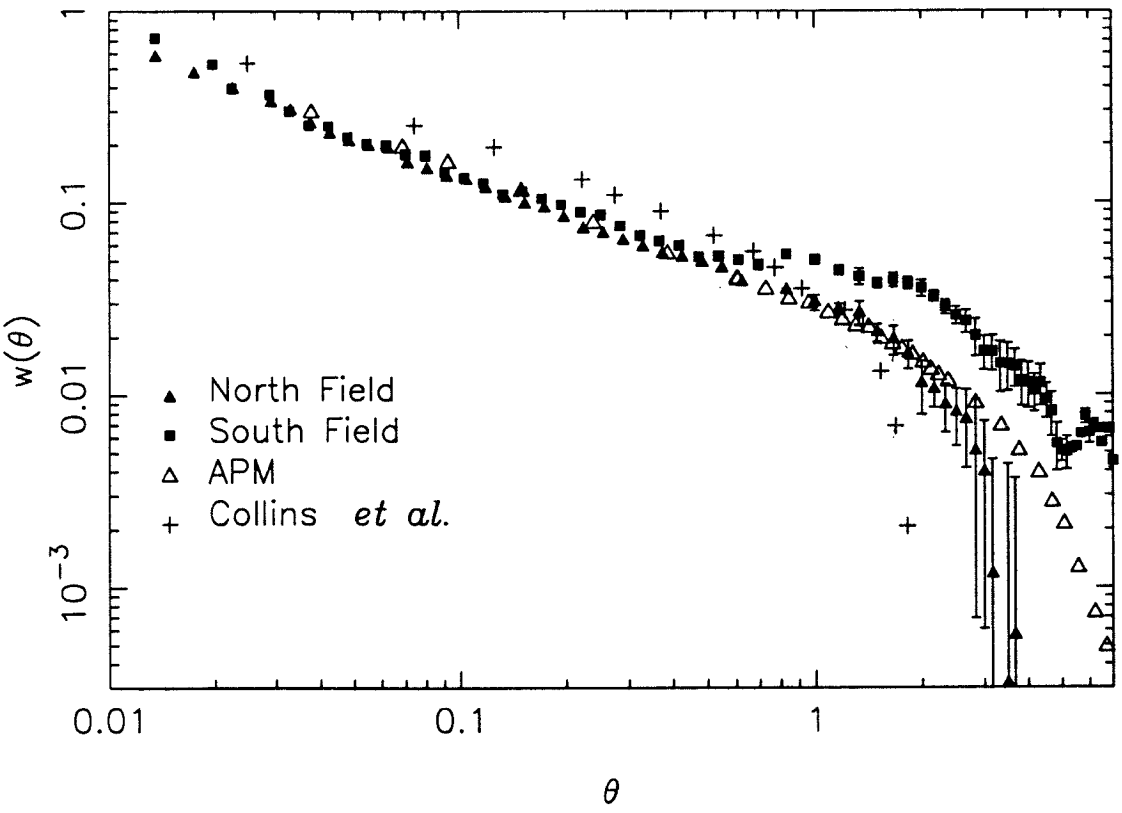
calibration available up to now, provides strong motivation to extend the present catalogue over the full area available in POSS-II.

I thank the group at COSMOS, and in particular H. T. MacGillivray for his support, both during and after the scanning of the plates. Also, many thanks are due to the Sky Survey team, for their expertise and effort in acquiring the plates for this work. I also wish to thank Michael Strauss for many useful and stimulating discussions. I also wish to thank the anonymous referee for helpful comments and suggestions.

3.5 Figure Caption

Figure 1. The two-point angular correlation function $w(\theta)$ as a function of angle in degrees. The filled triangles and squares are, respectively, our measurement for the north and south fields. The error bars come from a Monte-Carlo simulation in which artificial errors are introduced in each plate's magnitude scale and star-galaxy separation criterion. The open triangles, presented for reference, are the results of Maddox *et al.* (1989). No scaling was performed. The excellent agreement at small angles is evidence that the two surveys have comparable depths. The points of Collins *et al.* are shown by the crosses. The disagreement at small angles is easily understood as a result of the different depth of the samples, but the discrepancy at large angles is problematic.

Figure 1


w(θ)
θ
1
0.1
0.01
10^{-3}
0.01
0.1
1
North Field
South Field
APM
Collins *et al.*

3.6 References

Collins, C. A., Heydon-Dumbleton, N. H. and MacGillivray, H. T., 1989, *M.N.R.A.S.*, **236**, 7p.

Efstathiou, G., Ellis, R. S. and Peterson, B. A., 1988, *M.N.R.A.S.*, **232**, 431.

Geller, M. J., de Lapparent, V. and Kurtz, M. J., 1984, *Astrophys. J. (Letters)*, **287**, L55.

Groth, E. J. and Peebles, P. J. E., 1977, *Astrophys. J.*, **217**, 385. (**GP77**)

Kent, S. M., 1985, *Pub. Aston. Soc. Pac.*, **97**, 165.

MacGillivray, H. T. and Stobie, R. S., 1984, *Vistas in Astronomy*, **27**, 433.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989a, *M.N.R.A.S.*, **242**, 43p. (**MESL**)

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989b, *M.N.R.A.S.*, **243**, 692.

Peebles, P. J. E.,*The Large Scale Structure of the Universe*, Princeton University Press, 1980.

Picard, A., 1991, *Ph.D. thesis* in preparation.

Schneider, D. P., Gunn, J. E. and Hoessel, J. G., 1983, *Astrophys. J.*, **264**, 337.

Seldner, M., Seibers, B., Groth, E. J. and Peebles, P. J. E., 1977, *Astron. J.*, **82**, 249. (**SSGP**).

Shane, C. D., and Wirtanen, C. A., 1967, *Pub. Lick Obs.*, **22**, Part 1.

Thuan, T. X. and Gunn, J. E. G., 1976, *Pub. Aston. Soc. Pac.*, **88**, 543.

Valdes, 1982, *Instrumentation in Astronomy* IV, SPIE **331**,465.

APPENDIX

A.1 Estimators of $w(\theta)$

We have estimated the value of $w(\theta)$ using standard, well-known estimators. For small angles, we compute the distance between each possible pair of galaxies in the data, and make a histogram of the number of pairs found as a function of angular separation. We then compare this histogram to one computed from the pairs between the data and a random catalogue. This method is direct and straightforward; unfortunately it is extremely CPU intensive, as the computation increases as the square of N, the number of points. My code, which is 80% vectorized, takes nearly 18 CPU hours on a Convex to calculate the distance between each pair of galaxies in a sample with 70,000 galaxies. Since this had to be done a large number of times (to test various effects, as explained below), it was imperative to speed things up. Therefore, this direct approach was only used out to a distance of 0.7°. We only let the program search for neighbouring galaxies in a strip which is 1.4° wide (this is easy as the data are ordered by declination, so finding the index numbers of the galaxies to be searched is inexpensive. For this purpose, we use the routine HUNT (Press *et al.* 1986) In this way, we can roughly speed up the process by 15/1.4, a speedup of a factor of 10.

For larger angles, the data is binned into cells. Each cell is checked to insure that there are no deleted regions (due to bright stars *etc.*) that overlap with the cell, and that the cell is entirely contained inside the accepted regions of the catalogue. Contaminated cells are not used in this analysis. We use the remaining cells to determine $w(\theta)$ by Equation 2 given in Chapter 3. It is important to understand that this equation only provides an *approximation* to $w(\theta)$, and furthermore, the approximation is only valid for angles $\theta \gg \theta_{\rm bin}$. We tested that the results were not dependent on the binning angle $\theta_{\rm bin}$ for a reasonable range of angles. Of course, we wish to make $\theta_{\rm bin}$ as small as possible, so that this approximation at the joining angle (0.7°) will be accurate, but here again there is a tradeoff, as a small $\theta_{\rm bin}$ implies a large number of cells, and the computing time goes as the square of the number

of cells, *i.e.* as the *fourth* power of the reciprocal of $\theta_{\rm bin}$. In the end we decided that $\theta_{\rm bin} = 0.25^\circ$ was a good compromise.

Figure 1 shows the values of $w(\theta)$ calculated using each method for the south catalogue. The grid method systematically overestimates the value of $w(\theta)$. For small angles, this is easily understood, as we know that the grid method will not yield the correct answer for angles comparable to the grid size. The discrepancy for larger angles, however, which is of the order of 5×10^{-3}, is not so easy to explain. The grid size was varied, and it was found that the results are quite insensitive to $\theta_{\rm bin}$. We also moved the grid over by 1/2 of a grid size in each direction. The results are shown on Figure 2. The difference between the two estimates of $w(\theta)$ is of order 10^{-3}. This probably yields a good estimate of the noise due to gridding. The pairs method was also tested against a random catalogue, with the result that the largest deviation from zero was 5×10^{-4}, except for the first bin (where the approximation is not good, and has a value of 2×10^{-2}). It is reassuring that the grid and pairs method have similar shape (except at very large angles, where there is a bump on the curve of $w(\theta)$ for the grid method).

A.2 Obtaining Meaningful Error Bars on $w(\theta)$

In Chapter 3, we briefly discussed how we obtained error bars on $w(\theta)$. I will now present details of the simulation that was used to obtain these error bars.

First, it is important to note that the *statistical* error bars which come from the number of pairs of galaxies in any angular bin are extremely small, as there will be well over 500,000 pairs in any bin. Of course, the errors are not strictly $\sqrt{N}$ because of the non-Poisson nature of the problem, but even if the errors are three or four times as large as this, we still have a statistical error only of $\approx 10^{-3}$. In short, statistical errors do not play a significant role in the determination of $w(\theta)$ for large galaxy samples, though they are extremely important in the three-dimensional case, where the number of galaxies with redshifts is small, or in the angular correlation function of *clusters* of galaxies, where the number of objects is again small. And for large angles, where we use our binning method, we do not even have the possibility of using these statistical error bars, so some other method is in order.

The simulation to determine the sensitivity of $w(\theta)$ to possible systematic errors in the catalogue was performed as follows. We used the south catalogue, because it contains more plates. The method consists of introducing errors in the magnitude scale and star-galaxy separation algorithm which are realistic, and then comparing the measurements of $w(\theta)$ on this perturbed catalogue with the original measurement.

It is difficult to establish an error on the value of $\phi_{\rm cut}$, but let us do the following. We first calculate the mean and standard deviation for the 9 values of $\phi_{\rm cut}$ (one for each plate). The ϕ values, *by their very construction*, try to eliminate as much dependence on the parameters of each plate, such as seeing, magnitude limit, *etc.* because the width of the stellar ridge in the $\mathcal{S}$–mag diagram is normalized out (see Equation 4, Chapter 2). Simply put, this means that the stellar ridge in the ϕ–mag diagram should have the same width on each plate, and furthermore should be independent of magnitude. For this reason, we expect the $\phi_{\rm cut}$ to be the same for each plate, and the dispersion in these values is probably a good upper limit to the

error in the $\phi_{\rm cut}$ values. The values obtained are $rms(\phi_{\rm cut}) = 65$, with a mean $\phi_{\rm cut}$ of 2293. We then generated a set of random numbers which had a mean of zero and a variance of $3 \times 65 \approx 200$ and added these numbers to the original values of $\phi_{\rm cut}$, generating a new set of values $\phi'_{\rm cut}$. It is in this sense that the error bars are 3σ on the ϕ values.

At the same time, we also perturbed the magnitude zero points of each plate. This is easier, because we have a good estimate for the error in the zero points of the plates. This error consists of the total error in the CCD photometry (this is explained in detail in Chapter 2) combined with the error in determining the best fitting curve to the COSMOS magnitudes. This total error is $0.06^{\rm m}$. We drew numbers from a set of random numbers with a mean of zero, and a standard deviation of 0.06, and added these numbers to each object on each plate, and then truncated each plate to our nominal magnitude limit of $r = 19.0$. The complete parameters for the simulation are given in the table below:

Table 1. Parameters for Monte-Carlo Experiment

Field	$\phi_{\rm cut}$	$\phi'_{\rm cut}$	δ(m)
679	2336	2342	0.010
680	2179	2056	-0.050
681	2291	2219	-0.010
751	2257	2271	-0.060
752	2275	2446	0.070
753	2334	2210	-0.060
823	2230	2308	-0.030
824	2355	2251	-0.020
825	2380	2379	0.080

Note that this set in fact has a mean of –0.07 and standard deviation of 0.052. After reconstructing our galaxy catalogue from these perturbed plates, we computed the correlation function again, using our gridding technique. We did not recompute w(θ) using the pairs approach; this is expensive in CPU time, and unnecessary, because

the systematic errors being checked here will mostly affect only large angles; as for small angles ($\theta < 0.7°$) most pairs come from galaxies which are on the same plate.

Figure 3 shows the results for $w(\theta)$ in the original and perturbed south catalogues. We conclude from this experiment that the difference between the north and south fields is at a level of significance which cannot be attributed to systematic calibration errors.

A.3 The Effects of Geometry

An obvious difference in the two fields is their geometry; simply put, the south field is a nice square box, whereas the north field is an irregularly shaped area. When computing the correlation function, the dependence is removed by comparing to a random catalogue which has the same geometrical properties as the data catalogue, at least in the pairs algorithm. In the gridding algorithm, however, such dependence is not immediately apparent, as no random catalogue is involved. To see the effect of geometry on our measurement of $w(\theta)$, a subsample of the south (the 9 plate) field which had approximately the same geometry as the northern field was generated, and $w(\theta)$ for this new catalogue was recalculated. The results are shown in Figure 4. The very slight difference between the two curves convinces us that geometrical effects are well compensated for by our procedure for evaluating $w(\theta)$, and gives us great confidence that the difference between our north and south catalogues cannot be ascribed to geometrical differences alone.

A.4 The Effects of Varying Depths (Scaling Tests)

We can cut our catalogue in various non-overlapping magnitude slices, and then calculate the correlation function for each sub-catalogue. The luminosity function of galaxies is broad, but on the whole, each sub-catalogue samples galaxies with an effective depth (in h^{-1} Mpc) of $D = \mathrm{dex}(0.2(m - M_* + 25))$, where m is the magnitude of the slice, and we use a value of $M_* = -20.73$ (see § 4.5). In this case we expect the correlation function to scale as follows: the shallow catalogues will have a higher amplitude and a break at a larger angular scale than the deeper catalogues (see Peebles 1980). When we perform this test on our catalogue, we obtain a somewhat surprising result: this scaling law is obeyed reasonably well for our northern field (see Figure 5), but not at all for our southern field (see Figure 6).

This is quite surprising, so we performed some tests to see if one or more of the plates in the southern field was showing systematically different results from the others and could thus be blamed for this counter-intuitive behaviour. We calculated $\mathrm{w}(\theta)$ for each plate in the south (in our full magnitude range), and plotted the results on Figure 7. It is apparent that all the plates are exhibiting the behaviour.

Also, for each plate, we made our magnitude slices again, and produced a plot analogous to Figure 5 for each individual plate. Some examples are shown on Figure 8. Here the symbols are as follows: squares indicate galaxies in the range $16.5 \leq 17.0$, triangles those in the range $17.5 \leq 18.0$, and circles indicate galaxies in the range $18.5 \leq 19.0$. Here the scatter is quite large, as the sub-catalogues only contain a few thousand galaxies, but we can see that each plate more or less shows the behaviour exhibited by the south field as a whole.

This result remains puzzling, and reinforces our belief that the large scale structure in our two fields is systematically different. If galaxies truly are distributed mostly in sheets, filaments or "great walls", then perhaps in the southern field we are looking at one such two-dimensional structure edge on. In this case, probing to fainter magnitudes does not increase the effective depth of the survey in the expected manner, and we are simply probing deeper into the luminosity function of a

local structure. This might explain the lack of scaling. Whatever the explanation, the effect is real, and cannot be accounted for by any known systematic effect.

A.5 The Effect of Galactic Extinction on $w(\theta)$

In Chapter 3, we did not include any effects due to galactic extinction. The fields are at high galactic latitude ($|b| \approx 60°$), so the effect should be small. Nevertheless, for completeness, we ran a test in the south field, where we included an obscuration term $A_b = 0.1(\csc(b) - 1)$ (see Maddox *et al.* 1989). We plot the resulting and the original correlation functions in Figure 9. The results are virtually unchanged, and are within the error bars set by our previous Monte-Carlo experiment. We conclude that differential galactic extinction across the field is unimportant.

A.6 Star Counts and Intra-Plate Non-Uniformities

Not only is it crucial to calibrate the catalogue properly across several plates, but we must also be certain that there are no gross non-uniformities in the magnitude scales within a single plate. The first test concerning this was presented in § 2.2.5, where we considered the scatter in the calibration curves from the different CCD fields.

We can also attempt to estimate non-uniformities in the magnitude scale due to purely instrumental effects (such as vignetting, air leaks in the plate holder *etc.*) as follows. These effects are assumed to be the same on each plate, which allows us to collate the data from all our plates to estimate them. This is a method similar to that used by the APM group (Maddox *et al.* 1989) to estimate their vignetting effects.

For each plate, we construct a matrix of counts for the stars in the magnitude range $16.5 \leq r \leq 19.0$, which we denote C_{ij}^{k}. We used a grid size of 8 arc-minutes, which yields a matrix which is 40×40. When a cell is contaminated by a bad region, we flag its value with a large negative integer. We then stack the plates of either the north or south fields, averaging for each cell (i, j) the values from each plate that is not flagged, yielding a matrix M_{ij}, thus

$$\mathrm{M}_{ij} = \frac{1}{N} \sum_{k=1}^{N} \mathrm{C}_{ij}^{k} \qquad (1)$$

where for the south field, $N = 9$ plates. Figure 10 is a contour map of this stack for the south field. The solid contour is at the mean level of 14.5 stars per cell, the dashed and dot-dashed contours are 1.5 below and above this level. Notice that there is a slight systematic fall-off of the star counts towards the east edge of the plates; this indicates a systematic error in the magnitudes due to vignetting. We also show such a plot for galaxies in the south field in Figure 11.

We can try to estimate quantitatively the possible non-uniformities which are present as follows. First calculate the average of the values of M_{ij}, denoted $\overline{x}$, and

the standard deviation σ_o. For the south, using a binning angle of 8′, we obtain $\overline{x} = 14.56$ and $\sigma_o = 1.45$. Now, for strictly Poisson error, we would have expected a value for $\sigma_{Pois} = \sqrt{14.56}/\sqrt{9} = 1.27$ (where we used all 9 plates in the south). If we ascribe the difference between the expected and observed variances in the star counts entirely to an (uncorrelated) error in the magnitude scales (due to non-uniformity) σ_u, then we expect $\sigma_o^2 = \sigma_{Pois}^2 + \sigma_u^2$, which yields $\sigma_u = 0.70$. Then, using

$$\frac{\delta N}{N} = \frac{0.70}{14.56} = \ln(10)\ 0.6\ \delta m \tag{2}$$

where we have assumed that the star counts grow with the Euclidean slope of 0.6, we obtain $\delta m = 0.035^{\mathrm{m}}$. We show the result of this experiment for various binning angles in the table below:

Table 2. Estimated Intra-Plate Non-Uniformities

Bin Size(′)	Average	σ_{Pois}	σ_o	σ_u	$\delta(m)$
8	14.6	1.3	1.5	0.7	0.034
16	58.2	2.5	3.8	2.8	0.035
32	230.8	5.1	12.6	11.6	0.035
64	923.3	10.1	62.7	61.9	0.047

Of course, stars are not really distributed in a random manner. We show in Figures 12 and 13 the average stellar density plotted against galactic latitude. A strong gradient is noticeable, as there are fewer stars towards the galactic poles than near the plane. If we correct the values of our matrix M_{ij} for this gradient, using a model for star counts in the galaxy, or by simply fitting a straight line to Figure 12, for example (the results are not very sensitive to which method we adopt), we find a further reduction in our estimates of the plate non-uniformities (the $\delta(m)$'s tabulated in Table 2) of 10%. These non-uniformities can therefore cause problems in the determination of $w(\theta)$ by introducing spurious variations in the surface density of objects on the order of 3%, which is much smaller than the measured value on those scales.

A.7 Edge Effects

As mentioned in § 2.2.9, there is a discrepancy between the plate to plate magnitude transformation obtained from the overlap regions between the plates and the CCD calibrated magnitudes. The overlap regions yield residual errors in the zero points of the plates of and 0.07^{m} in the north and 0.10^{m} in the south, which are larger than predicted by the uncertainty in our CCD calibration procedure (0.06^{m}). Therefore we investigated possible systematic errors in the overlap region.

We used the matrix M_{ij} (shown as a contour plot on Figure 10) and collapsed it in both the X and Y directions (corresponding to east–west and north–south). Figures 14 and 15 show the values of $\sum_i M_{ij}$ and $\sum_j M_{ij}$ as a function of the free index. We also show the analogue of these figures for galaxies, in Figures 16 and 17. There the counting statistics are worse, and there is the danger that some large scale structures may not average out over our small number of plates. For example, Figure 14 shows the density of objects in a vertical strip as a function of horizontal distance on the plates. (All plates have North up and East to the left, therefore small values of X are towards the East in Figures 14 and 16, and towards the South in Figures 15 and Figure 17.) The worst deviation is clearly in Figure 14, where for $X < 5$, there are 50 stars missing from an average of 583, which, by the argument used in § A.6, Equation 2, corresponds to a magnitude difference of 0.07^{m}. This decrement is not seen in the corresponding figure for galaxies. This is further evidence that the magnitude transformation in the overlap regions should not be trusted, at least not in their raw (uncorrected) form, thus supporting the conclusion reached in § 2.2.9. This is why, rather than apply a necessarily uncertain correction to the magnitudes of a small number of galaxies, we have preferred to simply ignore the overlaps and calibrate the magnitudes entirely with the CCD data.

It is important to test for the effect of these regions on $w(\theta)$, however, so we generated a catalogue from the south catalogue in which the overlap regions were removed, leaving a catalogue which looked like a tic-tac-toe pattern. We measured $w(\theta)$ for this catalogue, and compared it to the original, as shown in Figure 18. The difference is largest at an angle of 4°, and the amplitude of this difference is 0.005.

This is measurable, but not substantial compared to the difference between $w(\theta)$ measured for the north and south fields. The power on scales of 2–2.5° is virtually unaffected, and cannot be reconciled with that measured in the north field, or with the results of Groth and Peebles (1977).

A.8 Correlation Function of Stars

We have computed the correlation function for objects classified as stars. At least over a fairly small angular range (where differential galactic extinction and stellar density variations due to Galactic structure are unimportant), we expect the stars to be randomly distributed. For a catalogue consisting of a single plate, this approximation seems to be valid. Figure 19 shows $w(\theta)$ for stars on Field 679. Note that the axes are linear, not logarithmic. There is a definite slope, but the amplitude is quite small; indeed $|(w(\theta))| < 5 \times 10^{-3}$ over most of the angular range.

For the complete survey, however, $w(\theta)$ is not so small. Figure 20 and 21 show $w(\theta)$ for all stars in the north and south fields. There we see quite a robust trend. Of course, this is not terribly surprising given the large (known) gradient of stellar density due to Galactic structure. We will show in the next chapter that in our catalogue, stars outnumber galaxies by a factor of 3–4, and that the stars are properly classified in 94% of the cases, yielding approximately 15% of the objects in the galaxy catalogue being in fact misclassified stars. Then the amplitude of $w(\theta)$ of 0.01 at 2° becomes a contribution of 0.00025 (if stars and galaxies are uncorrelated), which is much smaller than that observed, so the effect on the measured $w(\theta)$ for galaxies is unimportant.

A.9 Figure Captions

Figure 1. A comparison of the values of $w(\theta)$ computed using the direct counting of pairs and the gridding method. Note the discrepancy, especially at large angles.

Figure 2. An estimate of the error due to grid noise in estimating $w(\theta)$ using the grid method. Two estimates of $w(\theta)$ are shown, computed with the same grid size, but with one grid shifted by half a grid size with respect to the other.

Figure 3. Results for $w(\theta)$ for the original south galaxy catalogue, and the same catalogue with perturbed limiting magnitudes and $\phi_{\rm cut}$, as shown in Table 1.

Figure 4. The effects of the irregular geometry of the north catalogue on $w(\theta)$ were simulated on the south field. Such effects appear to be negligible.

Figure 5. We present $w(\theta)$ for the north catalogue in magnitude slices. Theory predicts that the fainter slices have a lower amplitude, and a break at smaller angular scale.

Figure 6. Same as in Figure 5, but for the south field. The scaling law is not obeyed.

Figure 7. The correlation functions of galaxies for each plate on the south field are shown superposed on top of one another. None of the plates exhibit any behaviour significantly different from the average.

Figure 8. We present $w(\theta)$ for some individual plate in the south field. We use this test to determine if the failure of the south field to scale properly with magnitude is due to one plate exhibiting a markedly different behaviour from the rest. This does not appear to be the case. The squares indicate galaxies in the range $16.5 \leq 17.0$, triangles those in the range $17.5 \leq 18.0$, and circles indicate galaxies in the range $18.5 \leq 19.0$.

Figure 9. The effect of differential galactic extinction of $w(\theta)$ is simulated by adding an appropriate error to the magnitudes of each galaxy. Such effects appear to be small in the region considered.

Figure 10. The variation of density of stars as a function of location on the photographic plate is shown. The data were obtained by stacking all 9 plates from the south field, and counting stars in squares 8′ on a side. North is up and east to the left. The solid contour represents a density of 14.5 stars per cell, the dashed and dot-dashed contours represent 1.5 stars per cell above and below this level. Note the systematic lack of stars on the east edge and north-east corner. This may be due to vignetting effects in the magnitudes.

Figure 11. Same as Figure 10, but for galaxies.

Figure 12. We plot the average number of stars in a cell as a function of the cells galactic latitude for all plates in the north field. (The cells are 8′ wide.)

Figure 13. Same as in Figure 12, but for the north field.

Figure 14. A plot of the mean surface density of stars as a function of their X position on the plate. This figure was obtained by collapsing the data used to produce Figure 10.

Figure 15. Same as for Figure 14, but for the Y direction.

Figure 16. A plot of the mean surface density of galaxies as a function of their X position on the plate. This figure was obtained by collapsing the data used to produce Figure 11.

Figure 17. Same as Figure 16, but for the Y direction.

Figure 18. We plot $w(\theta)$ computed from the south catalogue after strips had been removed from the edges of the plates. The differences between this catalogue and the original one are small, and cannot explain the difference between the north and south fields.

Figure 19. We show $w(\theta)$ for stars on Field 679. The amplitude is everywhere very small, in good agreement with the assumption that stars are randomly distributed.

Figure 20. The correlation function for all stars on the north field.

Figure 21. The correlation function for all stars on the south field.

Figure 1.

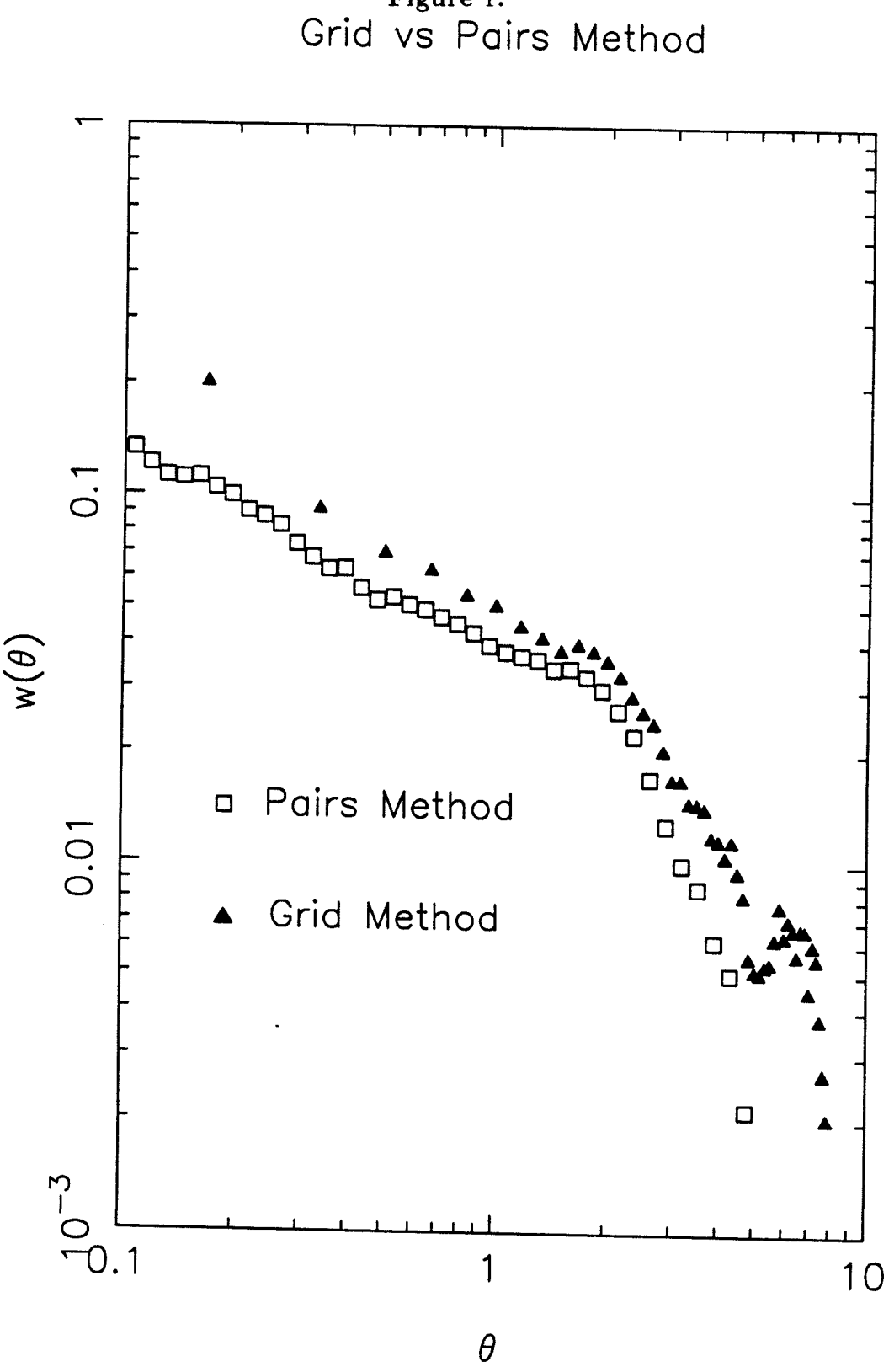

Grid vs Pairs Method
w(θ)
1
0.1
0.01
10^-3
0.1
1
10
θ
Pairs Method
Grid Method

Figure 2.

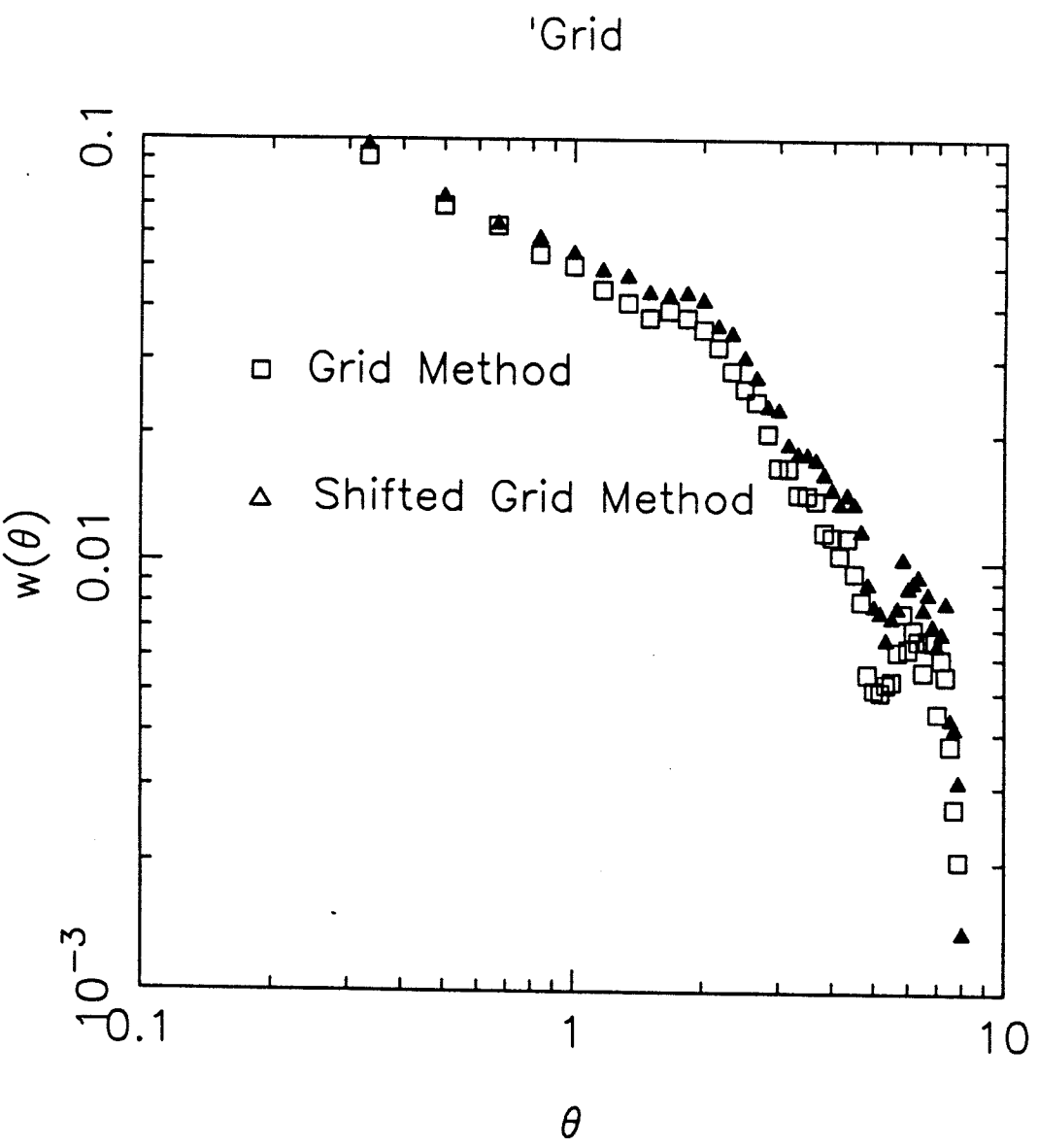


Figure 3.

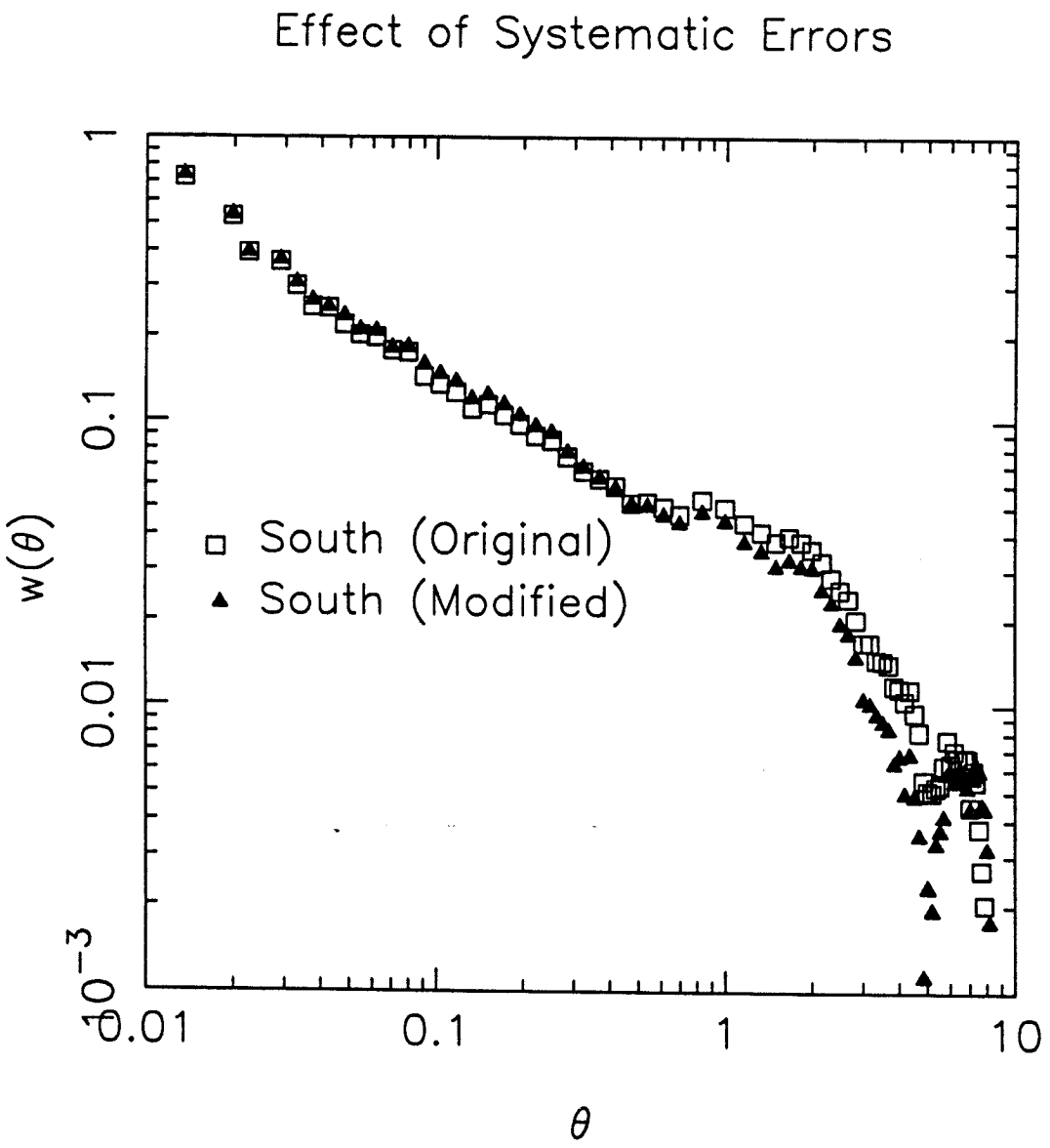

Effect of Systematic Errors
w(θ)
θ
South (Original)
South (Modified)
1
0.1
0.01
10^-3
0.01
0.1
1
10

Figure 4.

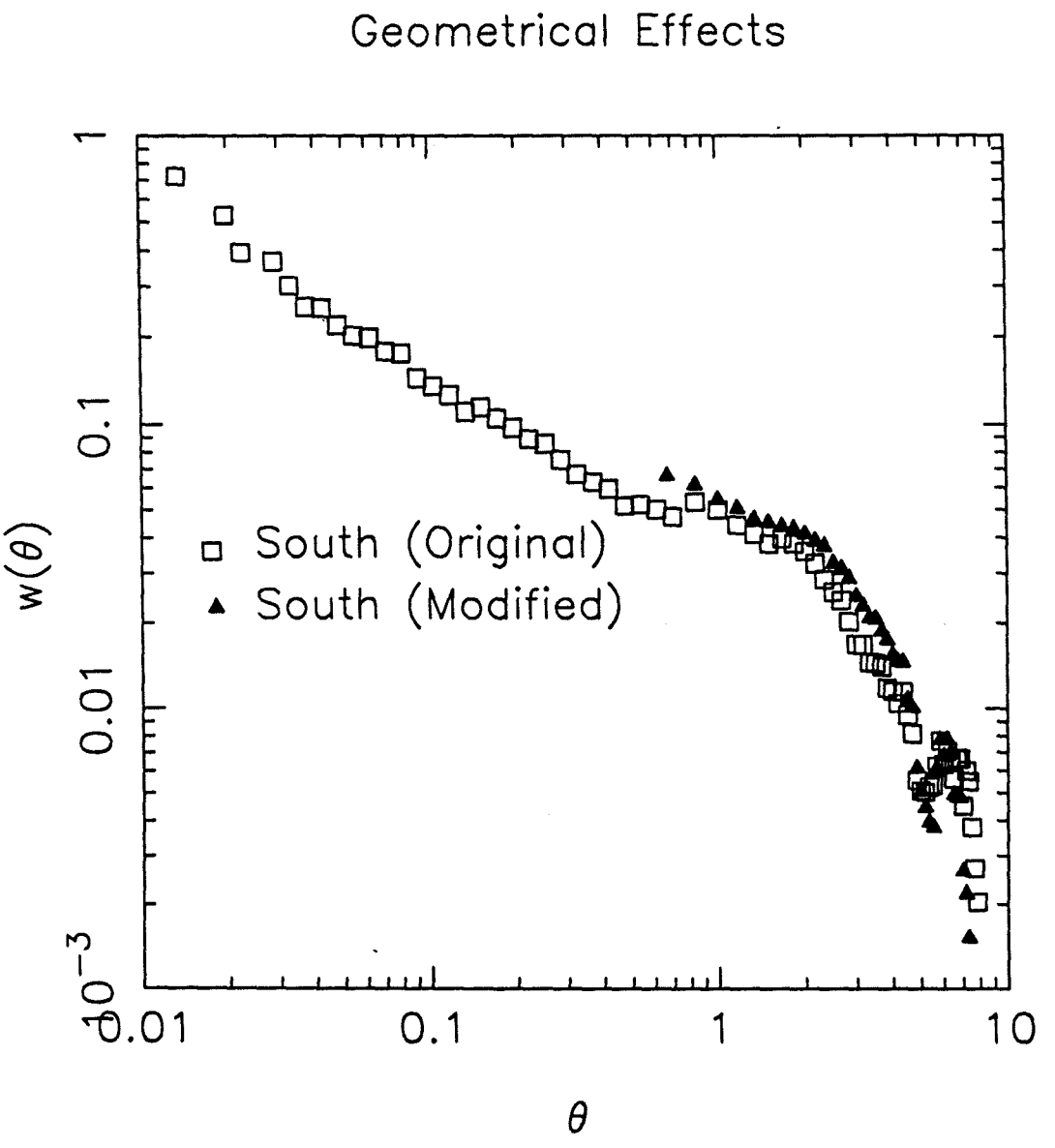

Geometrical Effects
w(θ)
θ
South (Original)
South (Modified)
1
0.1
0.01
10^-3
0.01
0.1
1
10

Figure 5.

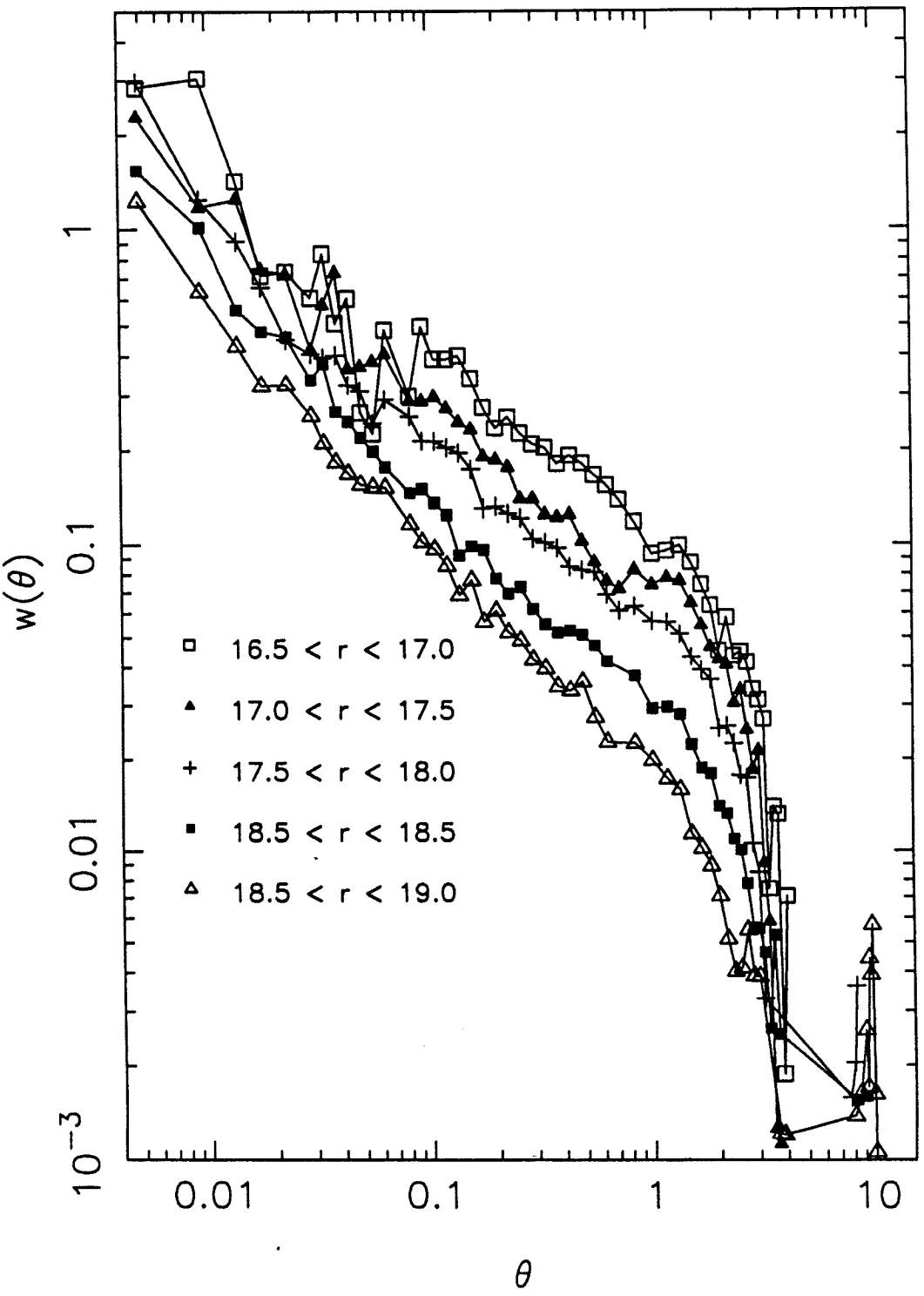

North Scaling
w(θ)
θ
1
0.1
0.01
10^-3
0.01
0.1
1
10
16.5 < r < 17.0
17.0 < r < 17.5
17.5 < r < 18.0
18.5 < r < 18.5
18.5 < r < 19.0

Figure 6.

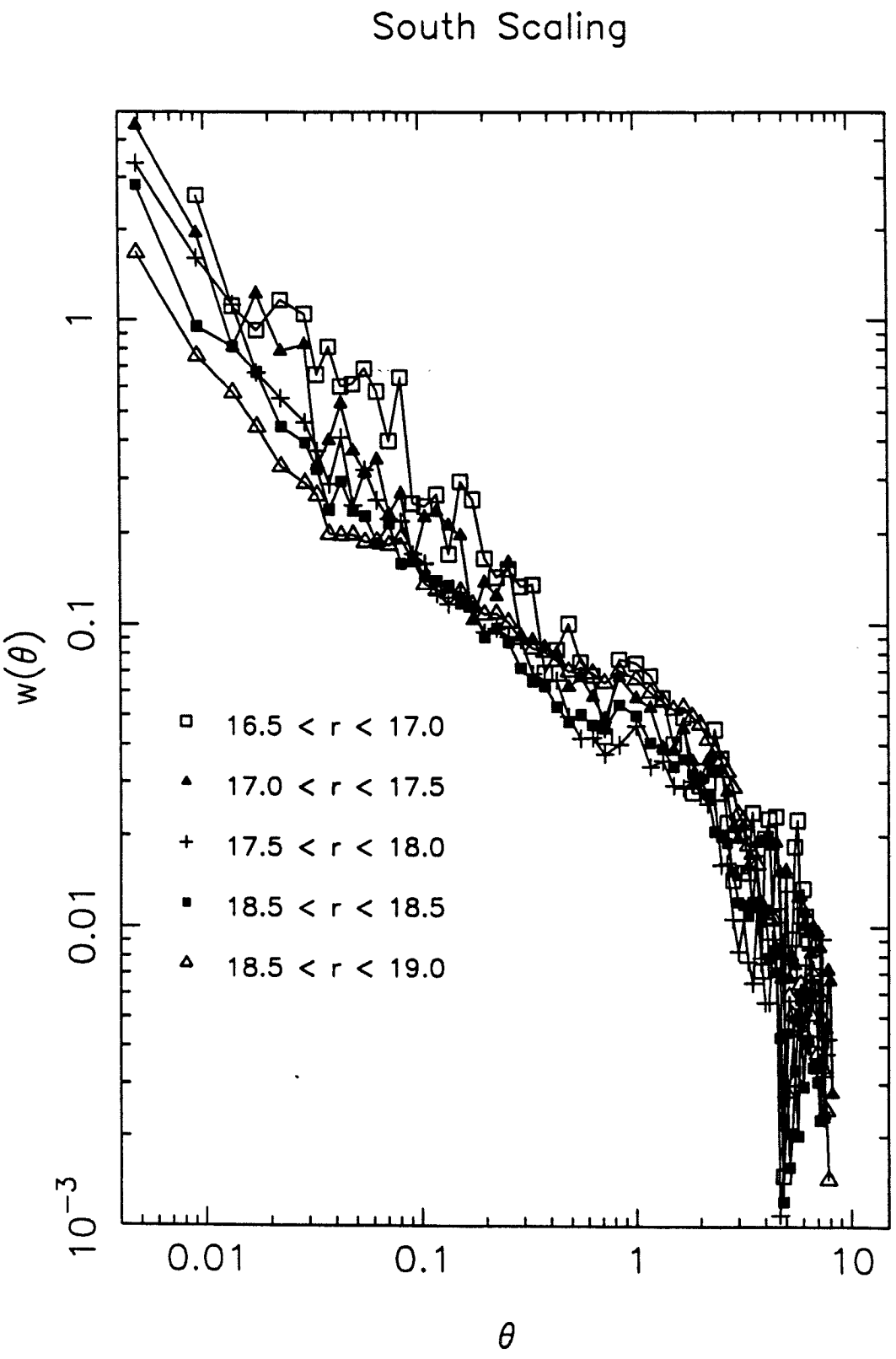

South Scaling
w(θ)
θ
1
0.1
0.01
10^-3
0.01
0.1
1
10
16.5 < r < 17.0
17.0 < r < 17.5
17.5 < r < 18.0
18.5 < r < 18.5
18.5 < r < 19.0

Figure 7.

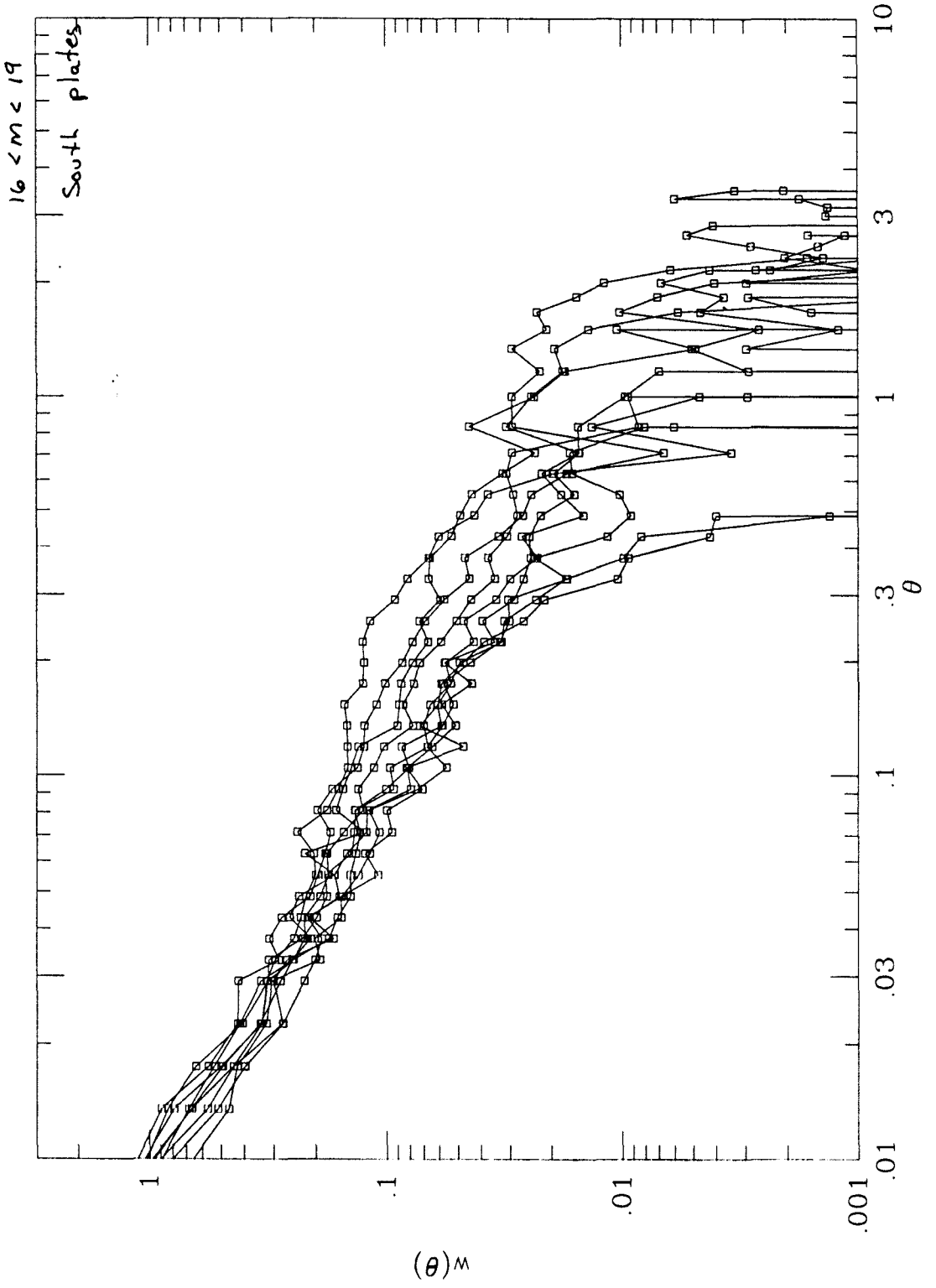


Figure 8.

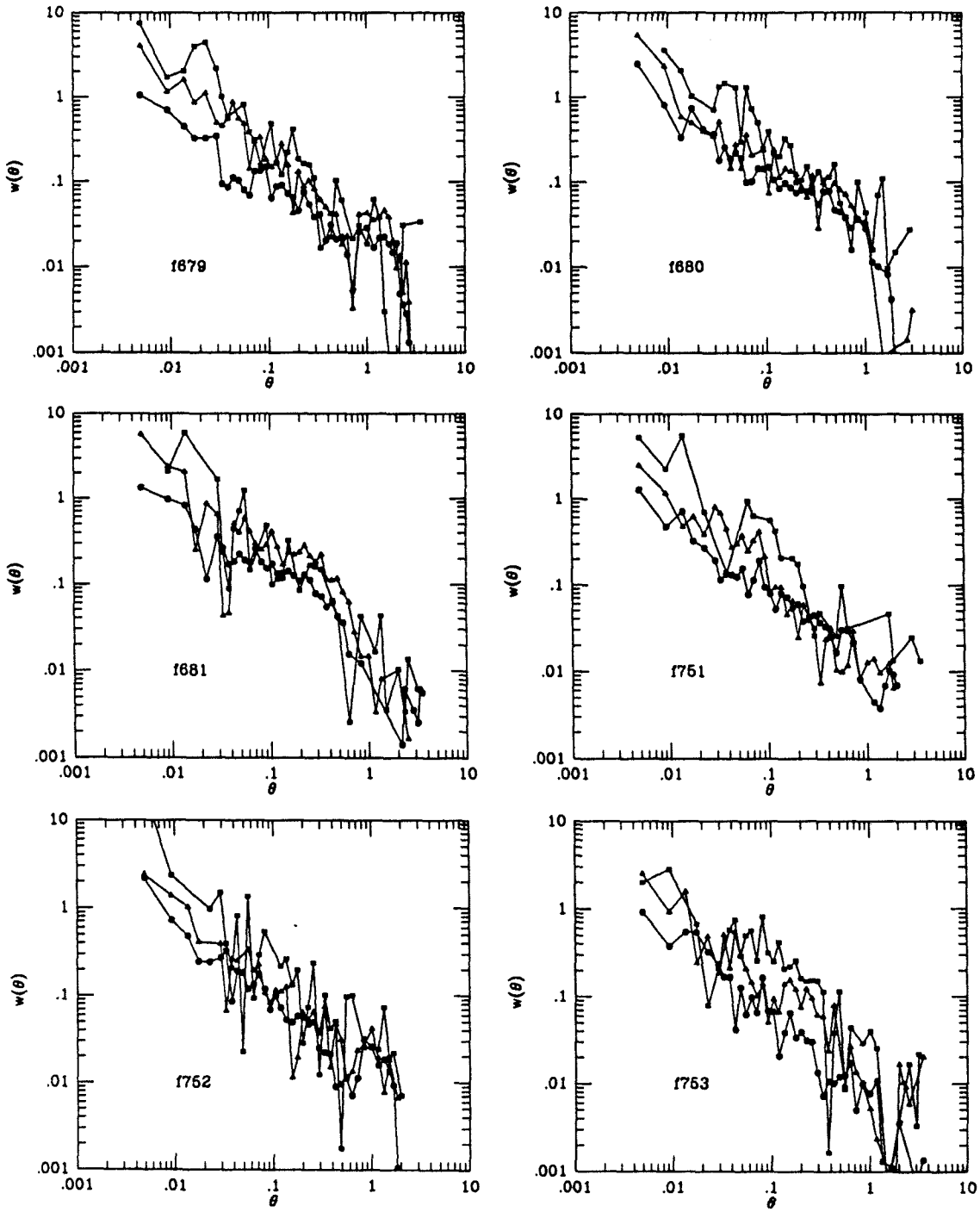


Figure 9.

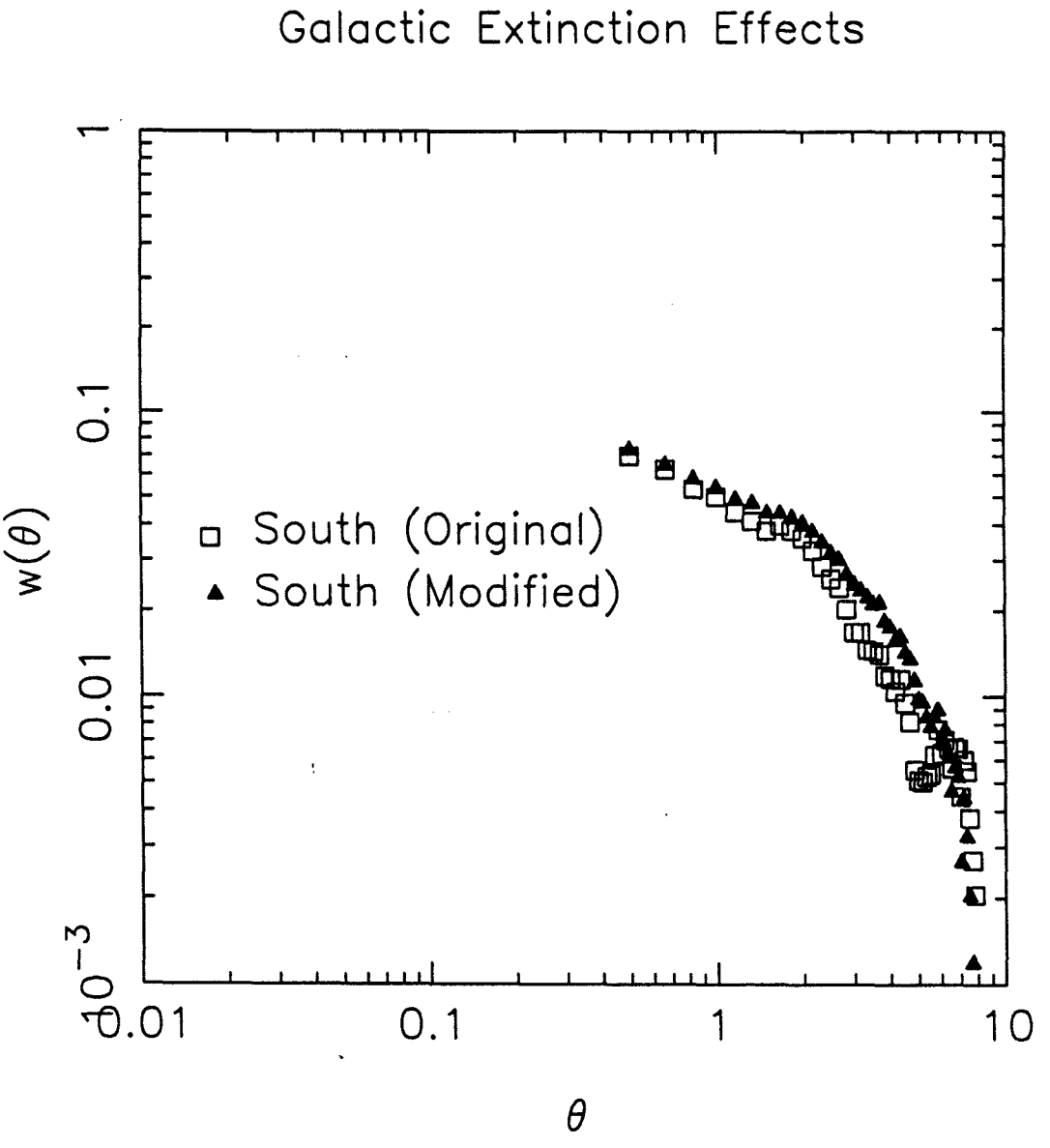

Galactic Extinction Effects
South (Original)
South (Modified)
w(θ)
θ
1
0.1
0.01
10^-3
0.01
0.1
1
10

Figure 10.

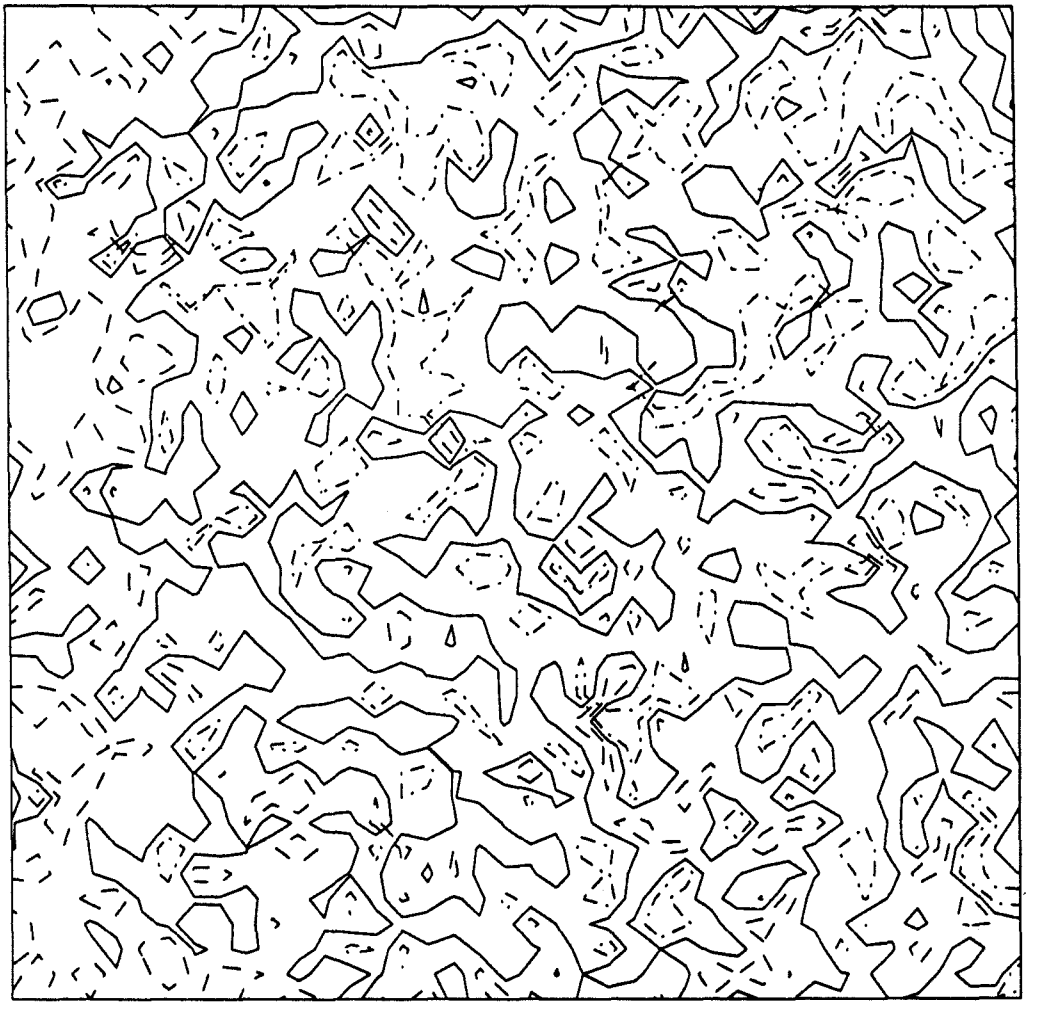


Figure 11.

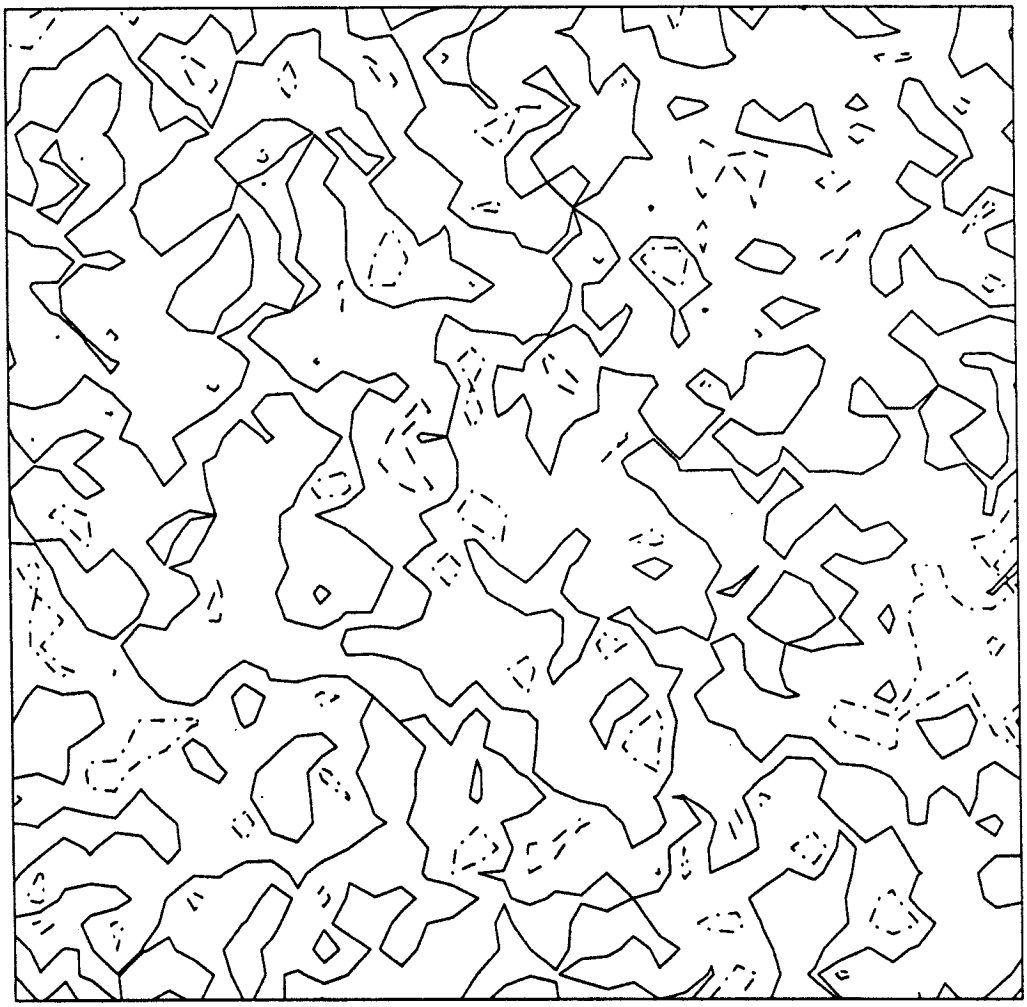


Figure 12.

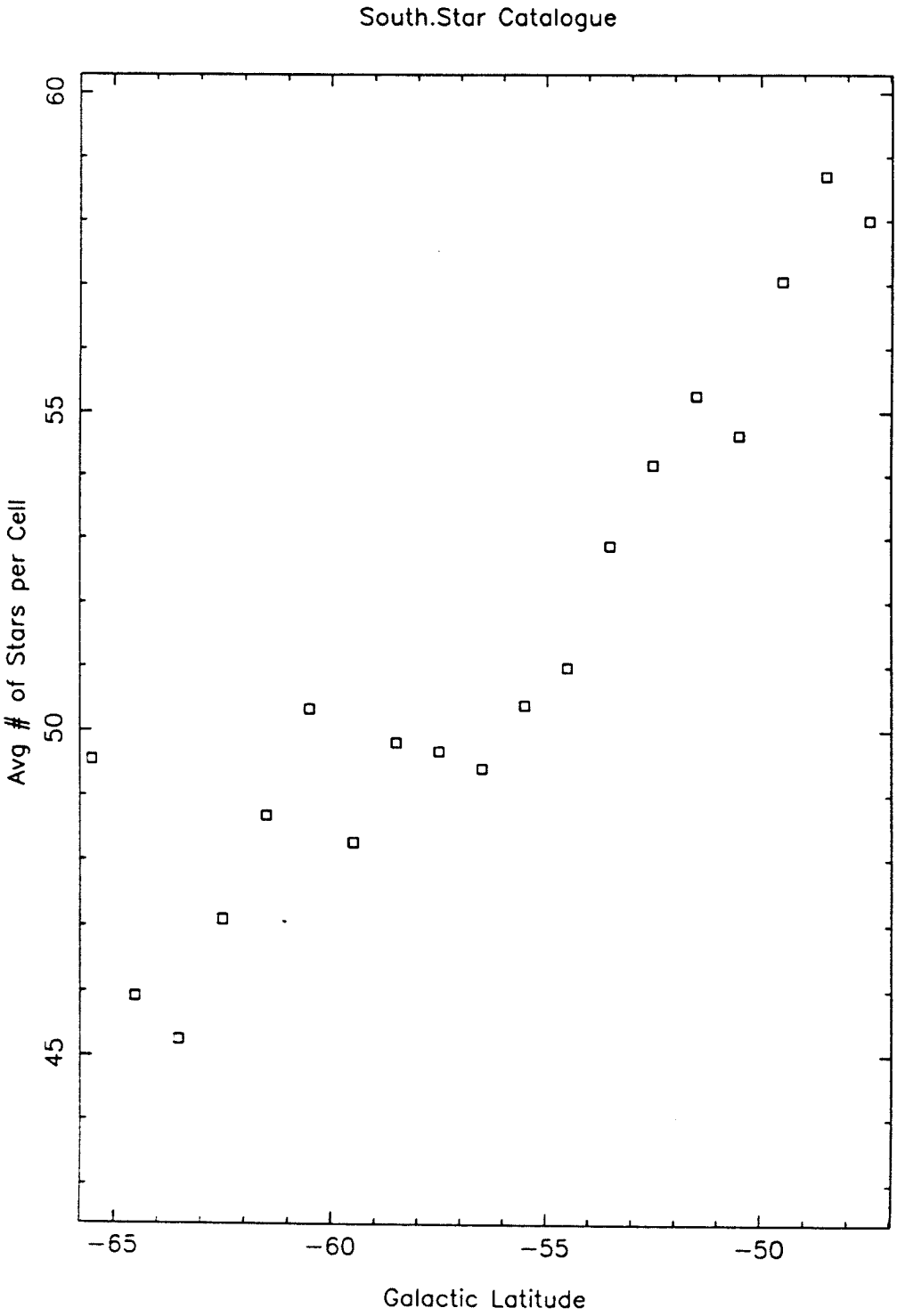

South.Star Catalogue
Avg # of Stars per Cell
60
55
50
45
−65
−60
−55
−50
Galactic Latitude

Figure 13.

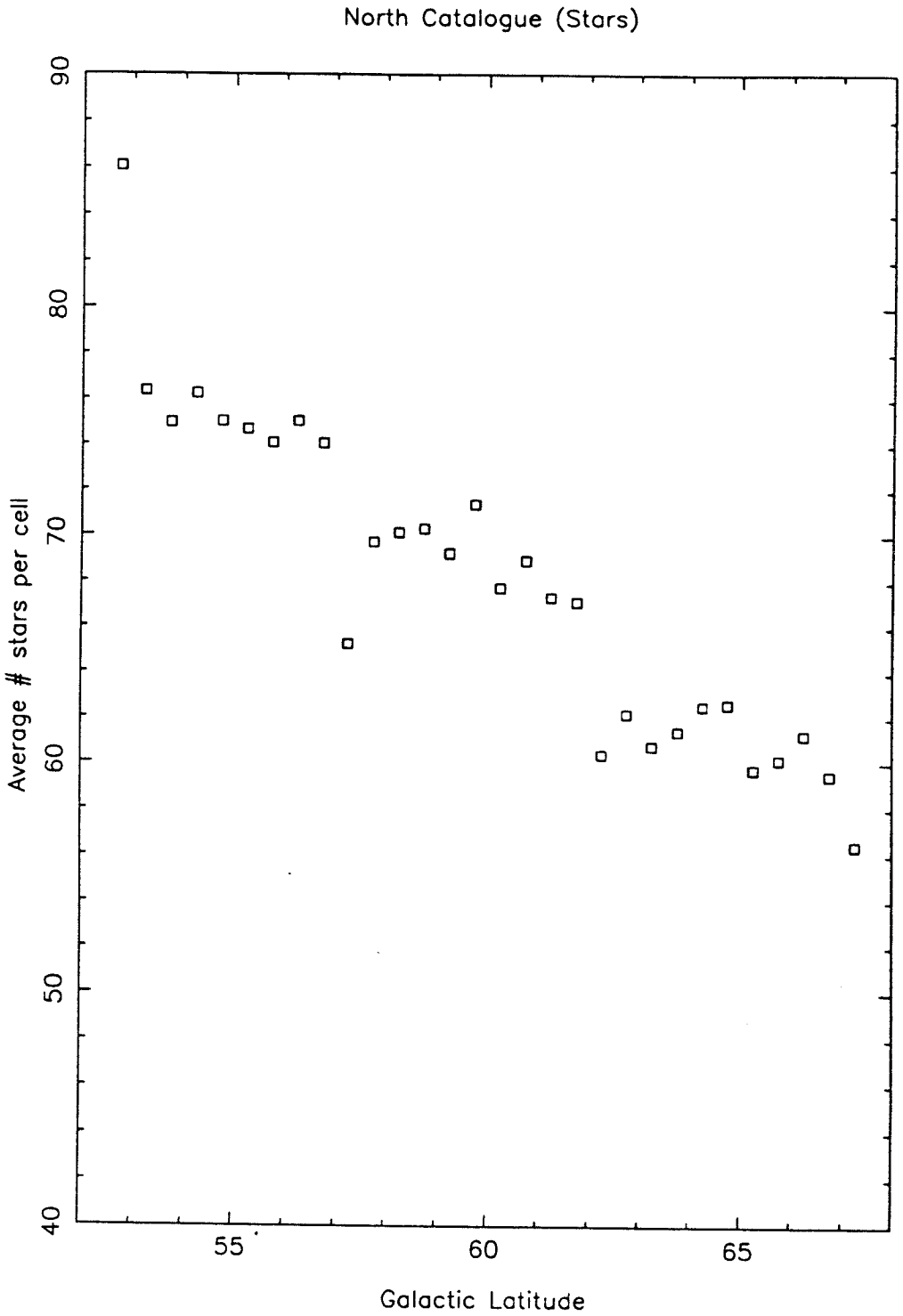

North Catalogue (Stars)
Average # stars per cell
90
80
70
60
50
40
55
60
65
Galactic Latitude

Figure 14.

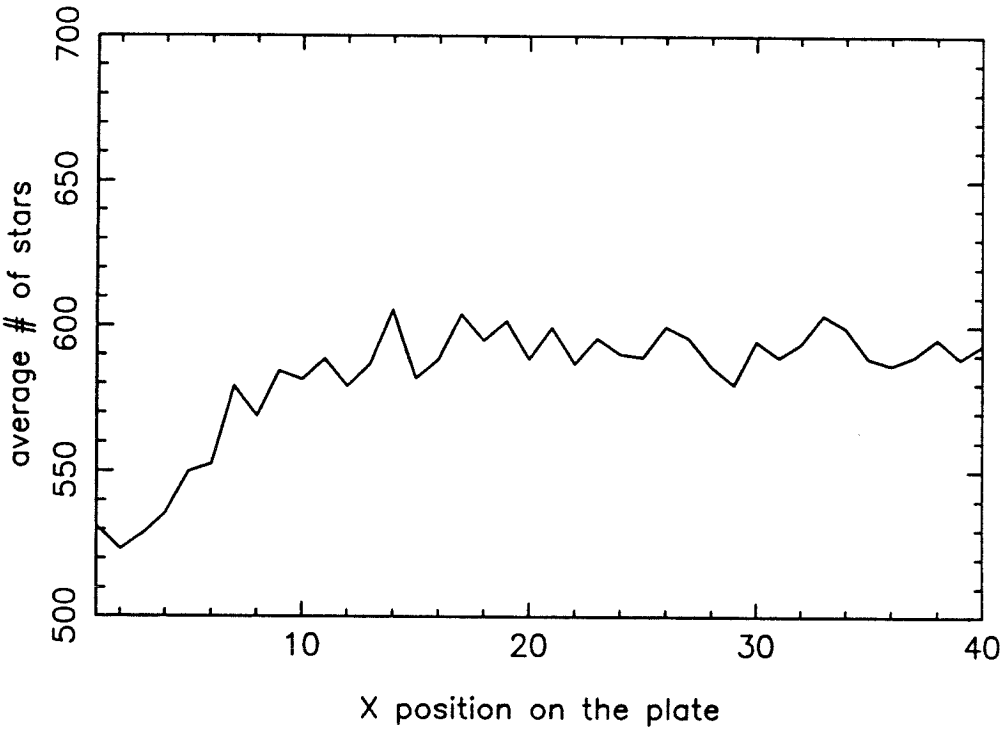


Figure 15.

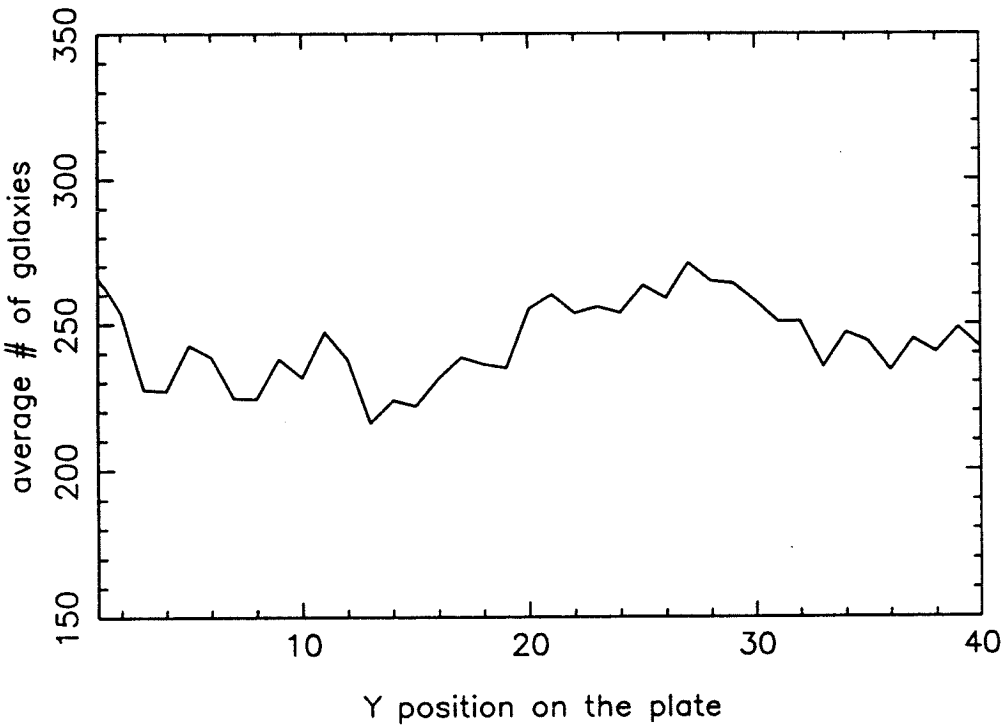


Figure 16.

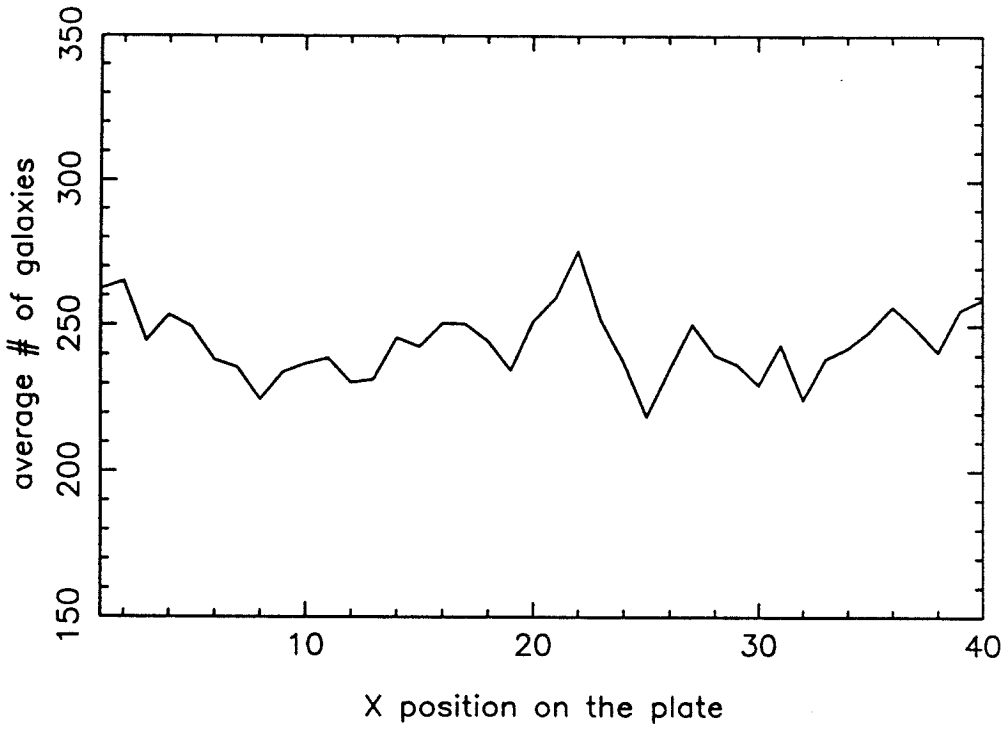


Figure 17.

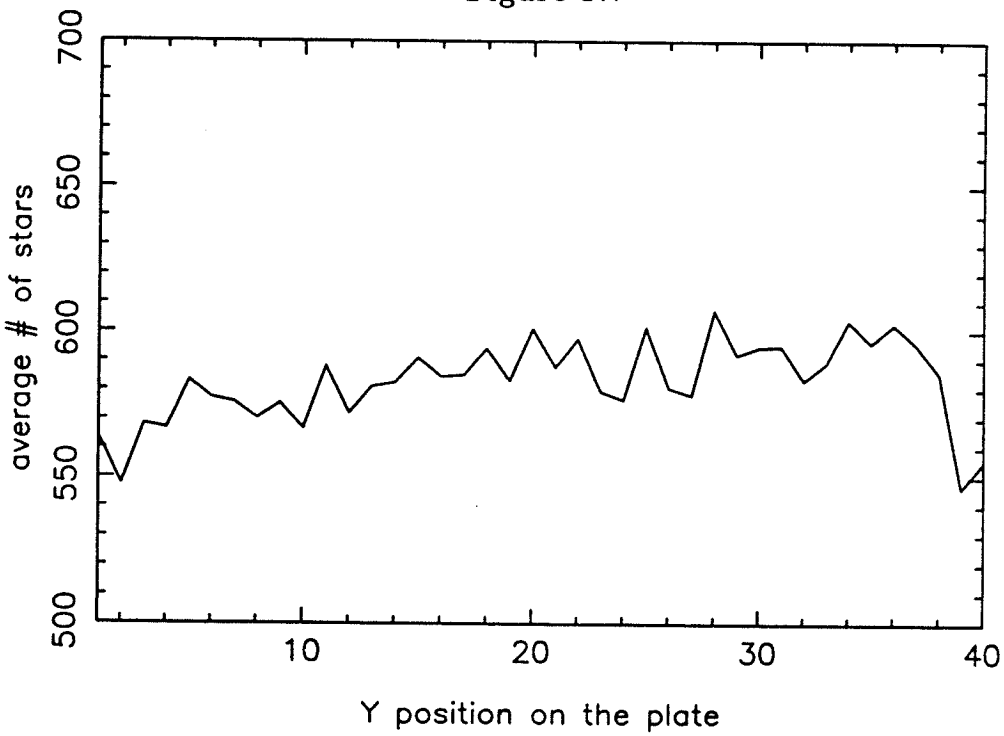


Figure 18.

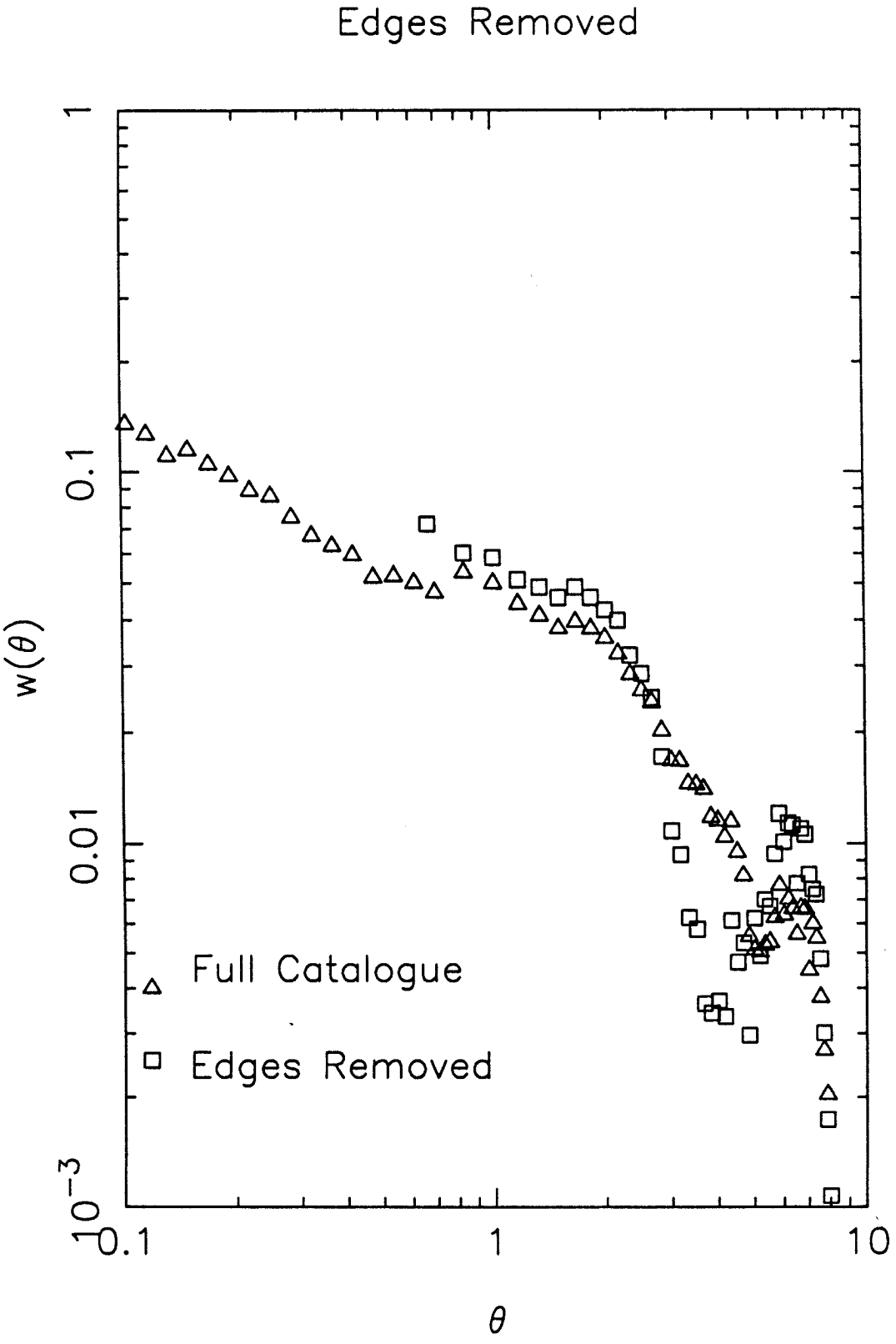

Edges Removed
1
0.1
0.01
10^{-3}
$w(\theta)$
0.1
1
10
θ
Full Catalogue
Edges Removed

Figure 19.

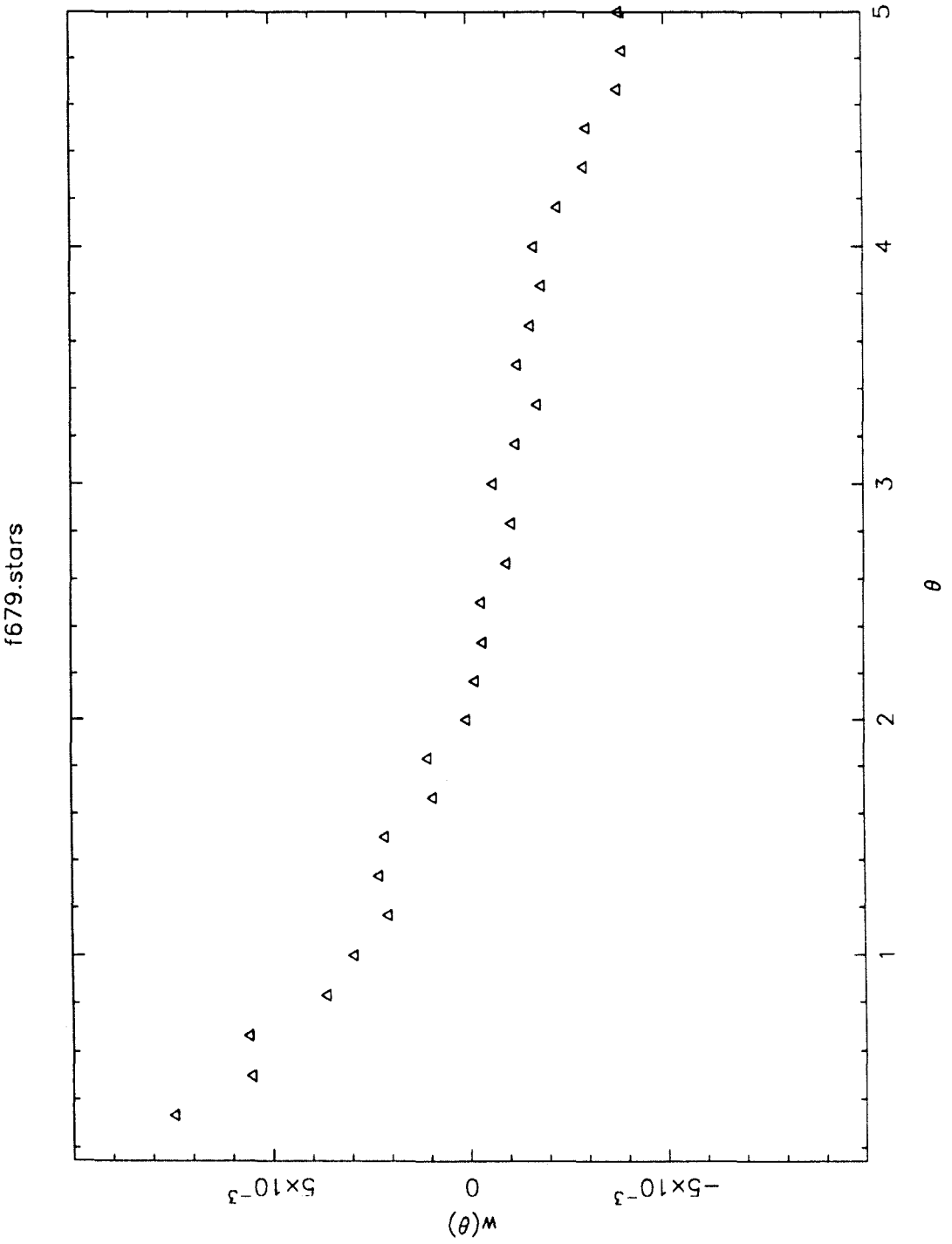


Figure 20.

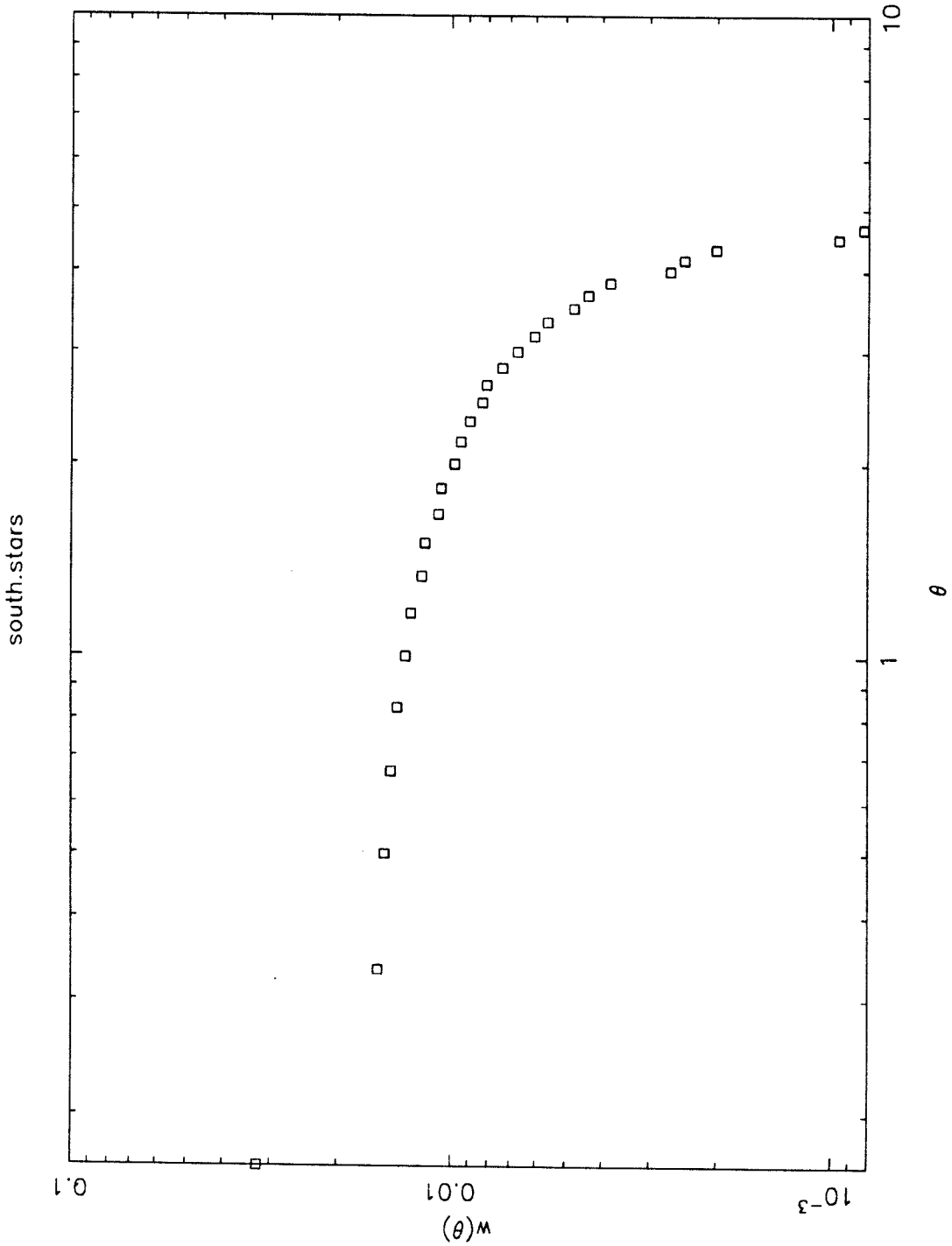
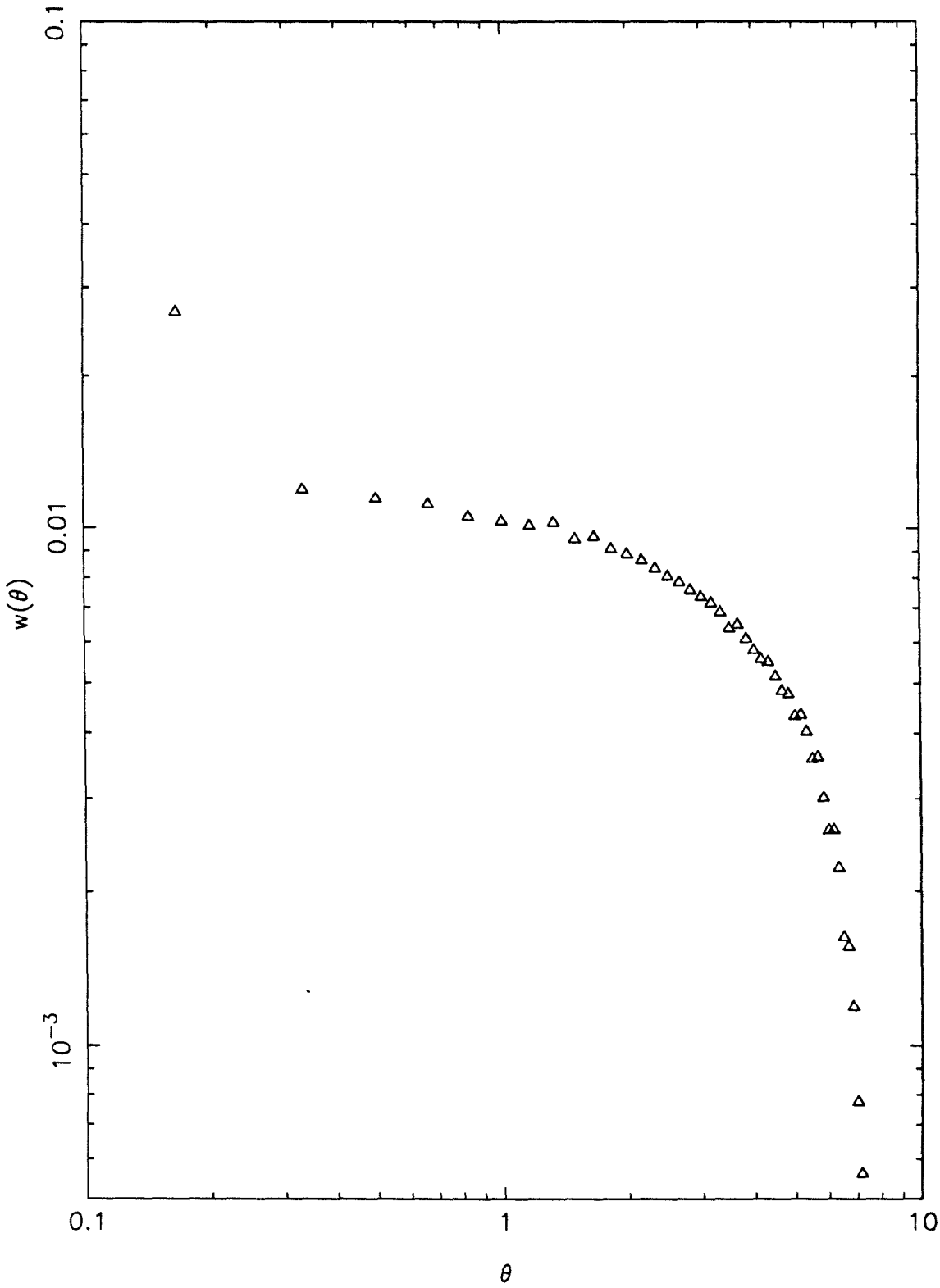


Figure 21.
North Stars



A.10 References

Groth, E. J. and Peebles, P. J. E., 1977, *Astrophys. J.*, **217**, 385.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989, *M.N.R.A.S.*, **242**, 43p.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. 1990, *M.N.R.A.S.*, **999**, 999. ref the photometry paper

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling., W. T., 1986, *Numerical Recipes*, Cambridge University Press

4. Inhomogeneities in the Universe† on Scales of $(125h^{-1}\,\mathrm{Mpc})^3$

4.1 Introduction

Very large scale structures in the distribution of galaxies such as giant voids and superclusters are no longer regarded as fanciful speculation. We now no longer doubt the existence of such structures, but rather seek to determine their dimensions, and how common they are. The distribution of galaxies is clearly non-random, but we are as yet unsure whether they lie on bubbles or sheets (de Lapparent, Geller, Huchra 1988), and whether this network is regular or random. The topology of these surfaces is unknown, as we do not know whether the voids or the sheets are the more fundamental entities (see Gott *et al.* 1986 for a discussion of sponge *vs.* meatball topology).

To address these issues with certainty will require samples which are considerably deeper than the structures in question. However, in every redshift survey conducted so far, structures have been found that are as large as could have been detected by the survey (Geller and Huchra 1988). The largest coherent structure yet found may extend to scales of $128h^{-1}$Mpc (Broadhurst *et al.* 1990), although more data are needed to confirm this result.

Several groups have reported number counts of galaxies extending over a wide range of magnitudes, mainly to study luminosity evolution of galaxies. Large fluctuations from field to field (in the 4 metre photographic surveys) and from author to author in deeper surveys (Tyson 1988, Shanks 1989, Koo 1981, Tyson and Jarvis 1981, Cowie *et al.* 1990, see Maddox *et al.* 1990 for a nice summary) have been reported. These fluctuations can be as large as a factor of two. There has been some controversy as to whether these fluctuations are due to real clustering or to differences in the magnitude zero point of the various surveys. The deepest surveys, such as Tyson's (1988) reach to $B = 27^m$ but cover only a few square arc minutes. Surveys at intermediate magnitude ranges, such as that of Koo (1981) still only

† This chapter is due for publication in the *Astronomical Journal* in August 1991.

cover about 1 square degree per field. It is possible that the discrepancies between these surveys was not mostly due to differences in photometric systems, or in data reduction procedures, but rather to real effects, *i.e.* , large scale structures.

Unfortunately, these surveys, except for that of Maddox *et al.* , only survey a modest volume of space. We have completed a galaxy survey down to $19^{\rm m}$ in r, in two widely separated regions of the sky. Each of these two fields probes a volume of size $\approx (125h^{-1}\,{\rm Mpc})^3$. The number counts for the galaxies in these two fields are different by 30%, with the north being the more populous field. Because all the plates have been brought to the same magnitude system via extensive CCD calibration, we cannot invoke field to field variations in the zero point of the magnitude scale to account for this effect: we would need a systematic shift in our magnitudes of $0.3^{\rm m}$ between the two fields to account for the galaxy counts.

Many other systematic errors are possible, however, and these are examined in the next sections. First, in § 4.2, we give a detailed description of the galaxy catalogue, the photometric calibrations that were applied, and describe the method used to perform star-galaxy separation and estimate the uncertainties of this method. Next, in § 4.3, we present tests which further convince us that we are probing to the same depth in both fields. These will include considerations of star counts, extinction, and the angular correlation function of galaxies $\mathrm{w}(\theta)$. Finally, in § 4.4, we present the results from our catalogue in the form of galaxy density contour maps and galaxy number counts. In § 4.5, a toy model is put forth to explain our observations.

4.2 The Galaxy Catalogue

4.2.1 General Description

Our data set consists of 16 IIIaF plates obtained at the 48" Oschin Schmidt telescope during the course of the second epoch Palomar Observatory Sky Survey (POSS-II). These plates were scanned at the COSMOS facility at the Royal Observatory of Edinburgh (MacGillivray and Stobie 1984). The COSMOS machine generates a set of parameters, including areas, magnitudes and peak intensities, for each image detected on the plate. These parameters are used to perform star-galaxy separation (see Picard 1991b[†]).

The plates are divided into two regions on the sky. The first of these consists of 7 plates, and is centered at $\alpha \approx 15^h, \delta \approx +30°, (l, b) \approx (45°, +61°)$. The second region consists of 9 plates and is centered at $\alpha \approx 0^h 20^m, \delta \approx +5°, (l, b) \approx (110°, -55°)$. We will henceforth denote these as our "north" and "south" fields respectively. Each plate has the central 5.3° scanned by COSMOS. The plate centers are offset by 5°, and the total area surveyed (after removal of bright stars, satellite trails etc.) is 386 square degrees.

4.2.2 Magnitude Calibration

For each plate, we obtained four CCD images of clusters of galaxies on the plate in the g and r bands of the Thuan-Gunn (1976) photometric system. This was done at the Palomar 60" telescope using a TI 800 × 800 CCD, and a re-imaging camera which yielded a pixel size of 0.62 arc seconds, and a total field of view of 8 arc minutes. The photometry resulting from these CCD frames is used to bring all 16 plates to a common magnitude system (Gunn r).

The CCD photometry was performed with the FOCAS (Tyson and Jarvis 1979, Valdes 1982) package, using isophotes of 24.61 r mag/["]^2 and 25.01 g mag/["]^2 for the isophotal magnitudes. These numbers were chosen because they represent 3% of the sky brightness at Palomar on a typical night. We used the FOCAS resolution

† This was a reference to this work, see Chapter 2

classifier to classify the objects on the CCD frames as stars or galaxies. We then compared the calibrated CCD magnitudes of all objects classified as galaxies by FOCAS to the raw COSMOS magnitudes of these objects. This is shown in Figure 1. We fit a second order polynomial (also shown in Figure 1) and applied this calibration to all objects on the plate. For objects brighter than $r = 16.5$, the calibration is not reliable, as the dynamic range of the emulsion exceeds that of the COSMOS machine, and there are too few bright objects to determine an accurate calibration. The *rms* value of the dispersion of the COSMOS magnitudes about the best fit in Figure 1 is 0.18^{m}, which is representative of the accuracy of the magnitudes measured from COSMOS. For objects fainter than $r = 19.0$, it is difficult to keep the star-galaxy separation uniform from plate to plate (see § 4.2.3), and so we limit our analysis to objects in the magnitude range $16.5 \leq r \leq 19.0$.

The CCD images were themselves calibrated by observing several standard stars each night. To achieve the greatest possible degree of uniformity in the magnitude zero points of each plate, only those nights on which the *rms* of the photometric solution was less than 0.035^{m} were accepted. Furthermore, the CCD frames used to calibrate any given plate were taken on different nights, and four calibration sequences were obtained for each plate. Thus the zero point error due to the calibration of the CCD photometry is $0.035/\sqrt{4} = 0.0175$. The formal error in the fit at $r = 19.0$ is 0.05^{m}, thus the total error in the zero point of any individual plate is 0.06^{m}.

The CCD photometry provides us with a test for inhomogeneities in the magnitude scale within a single plate. Figure 1 shows a typical calibration curve; the different symbols are for three different CCD frames (all lying on the same plate). A systematic deviation from one type of symbol over the others in Figure 1 would indicate a non-uniformity in the magnitude scale across the plate; no such non-uniformities are seen (see also Picard 1991*b* for other tests of magnitude non-uniformities within a single plate).

Table 1 shows details of the photographic material. The plate and field numbers, and date and time of acquisition are given, as well as the position of the center

of each plate. We have also included approximate galactic coordinates for quick reference. This table also shows the limiting magnitude of detection (in r) for each plate, which we define as the point where the differential number counts of all the objects on the plate (stars and galaxies) turn over.

4.2.3 Star Galaxy Separation

The procedure for separating stars from galaxies on the plates is described in detail in Picard (1991*b*), so we will only give a brief outline of it here.

The technique, as always, depends on the fact that stars, being unresolved, occupy a well determined locus in the parameter space measured by COSMOS. For example, at a fixed magnitude, all stars have approximately the same area, hence a tight *stellar ridge* is defined in a plot of area *vs.* magnitude for all objects. At a given magnitude, galaxies, because they are resolved, will have a larger area than the stars. This technique fails near the plate limit, where the two kinds of objects become indistinguishable, and in fact sets the limit on how faint we can keep the catalogue homogeneous. We use a parameter space defined by the magnitude, area, peak intensity and local sky intensity to perform the separation. We demand that this procedure be objective, so an algorithm is used to define which part of parameter space is to be assigned to stars, and which to galaxies, without the need for human intervention, to assure that we are following *exactly* the same procedure on each plate. The success of this procedure was then quantified by comparing the COSMOS classifications to visual classifications.

To obtain a statistically homogeneous catalogue of galaxies, therefore, it is not sufficient to impose stringent limits on the uniformity of the magnitude scale. We must also be assured that the probability with which a galaxy will be properly classified by COSMOS is uniform, *and* that the probability that a star will be properly classified is also uniform. We denote these quantities as η_g and η_s respectively. To this effect, we have examined randomly selected regions on each plate, and visually classified objects lying in the magnitude interval $16.5 \leq r \leq 19.0$ as stars or galaxies, and then compared these classifications to those of COSMOS to determine η_g and η_s. These visual classifications were done "blind", *i.e.* , we had no knowledge of

the COSMOS classification while visually classifying the images. This is extremely important, as otherwise, one's brain tends to learn what is a "success" and all objectivity is lost. The observer is trained to imitate the machine, rather than learn its performance.

The results are presented in Table 2. For each test region on the plate, we indicate how many stars and galaxies were visually classified (as N_{stars} and $N_{galaxies}$), and how many of each of these were properly identified by COSMOS, denoted by N^{ok}_{stars} and $N^{ok}_{galaxies}$. In the last columns we show the derived values of η_s and η_g. We note that Field 387 in our northern field has an unusually low value of η_s, and Field 679 in our southern field has a very low value for η_g. These two plates were flagged as being of inferior quality, and the global results for η_s and η_g for each field computed without taking these two plates into account. We estimate the uncertainties in these values by computing the root mean squared deviation from the mean, weighted by the number of objects in each field. We obtain $\eta_s = 93.5\% \pm 0.8\%$ and $\eta_g = 94.4\% \pm 2.2\%$ in the north, and $\eta_s = 94.2\% \pm 1.1\%$ and $\eta_g = 94.3\% \pm 1.5\%$ in the south. The values thus derived are fully consistent with the efficiency of star-galaxy separation being uniform between the north and south field.

4.3 Tests of Systematic Errors

In this section, we describe further indications that the large difference in the number counts is not due to systematic effects. We show that the star counts, and the angular correlation function of galaxies are not consistent with a difference in the zero point in the magnitude system of the two fields. We also examine how differential stellar contamination in the two fields accounts for some of the observed galaxy over-density in the north. Finally, the question of galactic extinction is considered.

4.3.1 Star Counts

To properly understand contamination effects, we need to examine star counts. Because our survey extends over large angles, effects due to our probing different regions of our Galaxy are to be expected; we can in turn use these to verify that there are no systematic errors in the galaxy catalogue.

Figures 2 and 3 show the differential star counts for each plate. The large discrepancy at bright magnitudes ($r < 16.5$) is due to poor calibration of the magnitudes. The sudden increase in slope at faint magnitudes is due to contamination by galaxies.

At any given magnitude, say $r = 18.0$, there is a large scatter in the differential number counts from plate to plate. We can see that this scatter is mostly due to our probing different regions of the Galaxy by comparing our observed star counts to those predicted by the model of Bahcall and Soneira (1980, using their Equation B1). The solid symbols in Figure 4 show the observed counts in $17.5 \leq r \leq 18.5$ compared with the counts predicted in this magnitude range at the central position of each plate by this Galactic model. The counts on the north plates are represented by squares, and those on the south by triangles. Since this model predicts counts in the V band, a shift in the ordinate to compensate for the difference in photometric system is justified. In other words, we are testing that the slope of the filled symbols will be close to unity (a formal fit yields 1.13). Clearly, most of the variance in the star counts can be explained by the position of the fields in the Galaxy.

Also shown in Figure 4 as the open squares are the observed star counts in the range $17.2 \leq r \leq 18.2$ for the north fields. These are plotted to test the hypothesis that there is an error in the magnitude zero point of 0.3^{m} in the north. The filled symbols agree with the Galactic model much better than the open squares and (filled) triangles. In other words, if we wish to reconcile the galaxy counts with a magnitude shift, we put ourselves in a position where our star counts do not trace the Galactic structure very well.

The agreement in Figure 4 is not perfect, but Bahcall and Soneira only claim 15% accuracy in the predicted counts. Also, their model assumes that the Galaxy is symmetric, and there is evidence that the sun is displaced slightly towards the north galactic hemisphere; see for example Figure 12 of Reid (1990), which shows that star counts in the south galactic pole are systematically higher (by about 10%) than in the north. If we were to displace the predicted star counts (the solid triangles) upwards by 10% on Figure 4, the agreement of the slope of the solid symbols with the model would be even better.

The agreement is nevertheless good enough to convince us that our counts are indeed tracing galactic structure, and thus provides an additional test that there are no large magnitude errors.

4.3.2 Differential Stellar Contamination

Since we have just seen that there are more stars in the north field than in the south, we can expect the north galaxy sample to have a higher proportion of stars masquerading as galaxies than the south sample. We can estimate the magnitude of this effect as follows. We first measure the observed ratio of stars to galaxies (denoted $\mathcal{R}_{\mathrm{obs}}$) in each field in our magnitude range of interest, and obtain $\mathcal{R}_{\mathrm{obs}} = 3.33$ in the north and $\mathcal{R}_{\mathrm{obs}} = 2.39$ in the south. Having measured the classification efficiencies η_g and η_s, it is a simple matter to solve for the *true* ratio of stars to galaxies, $\mathcal{R}_{\mathrm{true}}$. We obtain, thus,

$$\mathcal{R}_{\mathrm{true}} = \frac{\eta_g(1 + \mathcal{R}_{\mathrm{obs}}) - 1}{\eta_s(1 + \mathcal{R}_{\mathrm{obs}}) - \mathcal{R}_{\mathrm{obs}}}$$

which yields values of $\mathcal{R}_{\mathrm{true}} = 4.30$ in the north and $\mathcal{R}_{\mathrm{true}} = 2.73$ in the south.

Now, consider a population of 100 true galaxies and 273 true stars (in the south), and 100 true galaxies and 430 true stars (in the north). Then, given our η_g's and η_s's, we would *observe* $100 \times \eta_g + 273 \times (1 - \eta_s) = 110$ galaxies in the south and, by the same argument, 122 galaxies in the north. This corresponds to an $11\% \pm 4\%$ increase in the galaxy counts due solely to stellar contamination. (The uncertainty is computed using the error bars on η_s and η_g). It is thus not possible to account for all of the large over-density of galaxies in the north by appealing to differential stellar contamination.

4.3.3 Scaling of the Angular Correlation Function

Our last indication that the two fields probe to the same depth comes from the two point correlation function of galaxies $\mathrm{w}(\theta)$. The measurement of $\mathrm{w}(\theta)$ for these fields has been previously published by Picard (1991*a*). Here we use to our advantage the fact that there is a scaling relation for $\mathrm{w}(\theta)$ as a function of the depth (or limiting magnitude) surveyed (Limber 1954, Peebles 1980). We can calculate $\mathrm{w}(\theta)$ for any magnitude range. In particular, we can calculate it for galaxies in the range $16.2 \leq r \leq 18.7$, in order to simulate an error in the magnitude zero point. If such an error exists, then this new catalogue will actually sample galaxies to the same depth as the (original) south catalogue, and so $\mathrm{w}(\theta)$ will have the same amplitude at small angles.

Figure 5 shows the ratio of $\mathrm{w}(\theta)$ in the north field to the south field for angles between 0.02° and 0.5°. The filled squares are for both galaxy catalogues in $16.5 \leq r \leq 19.0$, and the open triangles for the *same* south catalogue, but where the north catalogue has been limited to the magnitude range $16.2 \leq r \leq 18.7$.

If the depths of the two samples are equal, we expect all the points to lie on $R = \mathrm{w}(\theta)_N / \mathrm{w}(\theta)_S = 1$. In fact, the south field exhibits stronger clustering at *all* angles (Picard 1991*a*), and so may in fact exhibit stronger intrinsic clustering than the north field. This might explain why R is below unity ($R \approx 0.9$). For the "bright" north catalogue, however, $R \approx 1.2$. We take this as further indication that our results are inconsistent with an error of 0.3^{m} in the zero point between the two fields.

4.3.4 Differential Extinction

The last possible systematic effect to be considered is differential galactic extinction through the two fields. We examined the reddening maps of Burstein and Heiles (1982), and found no emission above the lowest contour of $0.03^{\rm m}$ in B–V (or $0.1^{\rm m}$ in A_V). A difference in extinction of $0.1^{\rm m}$ between the two fields, with a slope for galaxy counts of 0.45 (where the slope α is given by $dN/dm \propto 10^{\alpha m}$) would lead to an apparent overdensity of 10%.

We have also used our CCD photometry to test for systematic differences in the galaxy colours. A difference in extinction between the two fields would result in a difference in the colour of a galaxy, so we can test this hypothesis by calculating the average colour for galaxies in each field, and assume that these are drawn from a representative set. For the north, we obtain, for the mean $g - r$ colour, $\overline{g-r} = 0.745$, with a dispersion of $0.43^{\rm m}$, calculated for 658 objects, whereas in the south, we obtain $\overline{g-r} = 0.768$, with a dispersion of $0.39^{\rm m}$, calculated for 477 objects. From this we obtain that the difference in the mean colours is 0.023 ± 0.025 magnitudes, which is consistent with no differential extinction.

4.4 Results

Maps of the surface density of galaxies in the north and south fields are shown in Figures 6 and 7. The data were binned in cells of 10 arc minutes, and then heavily smoothed to reveal small gradients in the general galaxy density field. The field of view is 15° on a side, East is up and North is to the left. The solid contour represents 400 galaxies per square degree (the global average for the 16 plates), and the dashed and dotted contours are the 20% above and below this value respectively. Note the presence of the Corona Borealis supercluster, located almost entirely on plate 449, which shows up as a large dashed contour near the eastern part of Figure 6. We wish to emphasize that the excess of galaxies in the north field is distributed over the *whole* area and is not due to one or two evident (small scale) structures. From this we will argue that whatever is causing the difference in density subtends an angle of at least 15°.

Differential number counts for both fields are shown in Figure 8. The data have not been corrected for incompleteness in detection of galaxies as a function of magnitude, but we expect this effect to be identical in both fields. For objects fainter than $r \geq 20.0$, the incompleteness becomes important, and the counts turn over. As mentioned in the previous section, the magnitudes are not well calibrated for $r < 16.5$, so we do not put much weight on those results. The slope of the counts, α, defined in the usual way where the differential number counts per magnitude interval $dN/dm \propto 10^{\alpha m}$, is 0.45 for the north and 0.46 for the south fields for objects in the magnitude range $16.5 \leq r \leq 19.0$. The striking difference in the counts between the two fields corresponds to a difference of 43% in the differential number counts of galaxies at $r = 18.4$. To explain this effect as a systematic error in the magnitude scales, we would need a shift in the zero point of 0.35^{m} between the two fields. We have shown in the previous sections how this is ruled out by our CCD photometry and other tests.

Figures 9 and 10 show the number counts for each individual plate in the north and south. For example, at $r = 18.4$, the mean counts and dispersion in the

north is 59.3 and 8.4 (6 plates) and 41.4 and 8.1 in the south (8 plates)[‡]. It is interesting to note that the counts are converging at $r = 19.0$ on the north field, but not on the south. The scatter from plate to plate is on the whole larger on the south fields. Note how on Figure 9, Field 449 clearly stands out as over-dense. As previously noted, the position of this field coincides with that of the Corona Borealis supercluster. This supercluster consists of the 6 rich Abell clusters A2061, A2065, A2067, A2079, A2089 and A2092, (Postman, Geller and Huchra 1986), which are *all* located on plate F449. Even with this plate taken out, however, the north field is substantially over-dense compared to the south field.

‡ Here we have omitted the two bad fields F387 and F679.

4.5 Discussion

Certainly, the most remarkable fact about Figure 8 is that not only is there a 40% difference in the number density of objects between the two fields, but that this difference is present over a range of $2\frac{1}{2}$ magnitudes!

We have examined all possible systematic errors, and find that we can explain 10% of the difference as being due to differential stellar contamination, leaving 30% to account for. Here we present a simple model to match the observations.

We first assumed a Schechter function (Schechter, 1976) with $M_* = -19.8$ in B_T (Efstathiou, Ellis and Peterson, 1988) and converting this value to our photometric system using the transformation given in Schneider *et al.* (1983). Schneider's transformation is for a brightest cluster galaxy with a $g - r = 0.47$, which is much redder than the typical galaxy. We adopt $g - r = 0.25$ and obtain $B_T = r + .93$. This yields a distance of $880h^{-1}$ Mpc for an L_* galaxy at $r = 19.0^m$.

In a Monte-Carlo experiment, we used this Schechter function to homogeneously distribute galaxies in metric coordinates in a cone of half-opening angle 7.5° and depth $1000h^{-1}$ Mpc. We used $q_0 = 0.5$ and the K correction suggested by Schneider *et al.* (1983). By adjusting the normalization of the luminosity function (ϕ_*), we can match the number counts quite accurately. The slope of the number counts in this model is 0.47, which is close to that observed (see § 4.4).

From this model, we can calculate the fraction of galaxies which lie within a given distance from us (see Figure 11). Conversely, we can use this relation to define a characteristic distance. We will define the distance which contains 80% of the galaxies in our sample as being the characteristic depth of the survey; from Figure 11, this is approximately $500h^{-1}$ Mpc. This distance is by necessity arbitrary, but the main point is that this sample certainly allows us to see large structures† to this distance.

The 16 plates are divided in two regions: 7 in the north galactic hemisphere, and

† We are currently acquiring redshifts for a complete sample of clusters of galaxies from this survey, and redshifts of $z = 0.25$ are not unusual.

9 in the south galactic hemisphere. The total volume surveyed is thus equivalent to a cube whose side has length $120h^{-1}$ Mpc in the north and $130h^{-1}$ Mpc in the south.

Maddox *et al.* (1989a) report that the *rms* fluctuations in cubes of side $150h^{-1}$ Mpc is 7%, based on their analysis of the angular correlation function of galaxies. The volume contained in each of our fields is comparable to this value. Our result, if due to chance, would then be two independent 2 sigma results (to explain away a 30% difference in galaxy density), and is therefore ruled out at the 99% confidence level. This indicates our sample and that of Maddox *et al.* are *not fair* samples of the universe.

If we make the reasonable assumption that the luminosity function is the same in both fields, then we are indeed measuring out to the same distance in the north and south. This means that this difference in the number counts is due to a real difference in the average density in these two fields.

What size of structure could be responsible for this? Because the luminosity function of galaxies is so broad, it is not difficult to come up with toy models to satisfy the constraint imposed by the number counts. For example, we have investigated putting down an over-dense spherical volume $200h^{-1}$ Mpc away (the distance to the Corona Borealis supercluster) on top of the smooth cone model mentioned above, and adjusting the overdensity of galaxies in this sphere to match the observed number counts. We find that we can obtain equally good agreement with small, very over-dense spheres as we can with larger, less over-dense spheres, for a large range of sphere radii. This is not surprising, for the number counts in a 2.5^{m} wide range alone are a rather poor constraint.

Infall velocities do provide some constraint, as we can predict the peculiar velocity which would result from such a structure using $\delta v/v = \frac{1}{3}(\delta\rho/\rho)\,\Omega^{0.6}$ (Peebles, 1980). Here $\delta\rho/\rho$ is the overdensity within a sphere of radius $200h^{-1}$ Mpc, and is thus related to the overdensity of the sphere ϵ of radius R by $\delta\rho/\rho = \epsilon(R)\,(R/200h^{-1}\,\mathrm{Mpc})^3$. So we pick a value R, calculate what ϵ is needed to fit the number counts, and then predict a local infall velocity towards this over-dense

sphere. In an $\Omega = 1$ universe, a sphere of radius $80h^{-1}$ Mpc requires $\epsilon = 1.13$, and would thus cause a local infall velocity of 480 km/s towards the sphere. Since the velocity of the Local Group with respect to the microwave background is 622 km/s (Smoot *et al.* 1991), we can not make the sphere much bigger than this or we will start predicting velocities larger than observed.

Our second constraint is that, as noted previously, there is no obvious structure in the two-dimensional maps of galaxies (see Figure 6) which can explain the excess of galaxies in the north, so the structure responsible for this excess must subtend an angle at least as large as the observed solid angle. If it is located at the distance of the Corona Borealis supercluster, this implies a linear extent of least $50h^{-1}$ Mpc. For a sphere of this radius, our toy model would require an overdensity $\epsilon \approx 2$ to account for the number counts.

Of course, the north could be the "normal" region, and the south under-dense. Our toy model works equally well with spherical voids in the south as with spherical overdensities in the north. In this case, however, the void needs to be large enough so as to make $\epsilon \geq -1$, which corresponds to a radius of $80h^{-1}$ Mpc. This void could therefore be comparable to the one in Boötes found by Kirshner *et al.* (1981).

In the standard model where large scale structures are created by gravitational instability without re-heating, our toy model would predict fluctuations in the microwave background radiation of order $\delta T/T \approx 2 \times 10^{-4}$ on 1° scale in an $\Omega = 1$ universe with biasing of unity (Juszkiewicz *et al.* 1987). This is well above the limits of non-detection of 2.7×10^{-5} (Meinhold and Lubin, 1991), and so we must conclude that either that such overdensities are indeed very rare (*i.e.* , this is indeed a one in a hundred event, as suggested by the result of Maddox *et al.*), or that the conditions of this standard model do not apply.

I thank the group at COSMOS, and in particular H. T. MacGillivray for his support, both during and after the scanning of the plates. Also, many thanks are due to the Sky Survey team, for their expertise and effort in acquiring the plates for this work. I thank Michael Strauss and Helen Johnston for their critical readings of the manuscript, and the latter for helping me enter my visual classifications in

computer readable files. Finally, I also wish to thank Sandra Faber for encouraging me to publish this result.

Table 1. Photographic Data

Plate Number	Field Number	Date	UT	α , δ (1950.0) (center)	(l, b)		Limiting r Magnitude
SF01814	386	11-May-88	6:23	14 49 34 +34 50 25	56.7	63.3	21.1
SF01109	387	29-Mar-87	0:00	15 13 28 +34 51 40	55.7	58.4	20.7
SF01098	447	27-Mar-87	8:58	14 35 36 +29 50 15	45.3	66.5	20.5
SF01824	448	12-May-88	6:40	14 58 38 +29 50 45	45.7	61.5	21.0
SF01165	449	21-Apr-87	7:48	15 21 32 +29 52 00	46.5	56.6	21.2
SF01094	513	26-Mar-87	0:35	15 03 38 +24 51 15	35.9	59.8	20.8
SF01183	514	27-Apr-87	8:45	15 25 36 +24 52 10	37.9	54.9	21.4
SF00952	679	27-Nov-86	3:44	00 01 53 +10 12 17	104.5	-50.7	20.4
SF02249	680	12-Dec-88	2:20	00 21 59 +10 12 50	112.2	-51.8	20.6
SF00973	681	3-Dec-86	2:52	00 41 55 +10 12 05	120.1	-52.4	21.0
SF02242	751	11-Dec-88	3:00	00 01 59 +05 12 53	102.0	-55.5	20.6
SF00743	752	1-Sep-86	8:35	00 21 54 +05 12 10	110.6	-56.7	21.6
SF01462	753	19-Sep-87	8:23	00 41 57 +05 12 20	119.7	-57.3	21.1
SF02251	823	13-Dec-88	3:00	00 02 00 +00 13 00	98.8	-60.2	20.6
SF01376	824	23-Aug-87	9:10	00 21 56 +00 12 30	108.6	-61.6	20.9
SF01478	825	24-Sep-87	8:15	00 41 56 +00 12 20	119.2	-62.3	21.2

Table 2. Visual Inspections

	N_{stars}	N^{ok}_{stars}	$N_{galaxies}$	$N^{ok}_{galaxies}$	η_s	η_g	
region1.387	121	95	45	43	78.5	95.6	
region2.387	85	63	34	33	74.1	97.1	
region1.386	112	107	63	61	95.5	96.8	⇐ North
region2.386	116	104	51	51	89.7	100.0	Fields
region1.447	168	156	79	76	92.9	96.2	
region1.448	147	137	50	46	93.2	92.0	
region2.448	84	78	39	39	92.9	100.0	
region1.449	105	96	105	89	91.4	84.8	
region1.513	151	146	35	35	96.7	100.0	
region1.514	140	133	27	27	95.0	100.0	
north(total)	1229	1115	528	500	90.7	94.7	
north(total)	1023	957	449	424	93.5	94.4	F387 Omitted
region1.679	51	50	31	15	98.0	48.4	
region2.679	66	62	38	24	93.9	63.2	
region1.680	144	129	38	35	89.6	92.1	⇐ South
region2.680	64	64	19	17	100.0	89.5	Fields
region1.681	87	74	23	22	85.1	95.7	
region2.681	114	111	23	23	97.4	100.0	
region1.751	44	41	30	28	93.2	93.3	
region2.751	84	78	29	26	92.9	89.7	
region1.752	104	100	68	65	96.2	95.6	
region1.753	90	87	28	28	96.7	100.0	
region2.753	68	66	22	19	97.1	86.4	
region1.823	83	78	37	36	94.0	97.3	
region2.823	94	92	22	17	97.9	77.3	
region1.824	66	62	37	35	93.9	94.6	
region2.824	79	71	43	42	89.9	97.7	
region1.825	105	102	38	38	97.1	100.0	
south(total)	1343	1267	526	470	94.3	89.4	
south(total)	1226	1155	457	431	94.2	94.3	F679 Omitted

4.6 Figure Captions

Figure 1. Typical calibration curve. Shown are the calibrated r magnitudes for objects classified as galaxies by FOCAS versus the raw COSMOS magnitudes. The relationship is highly non-linear; this is due to lack of proper density to intensity calibration during the scanning phase. Each CCD frame is consistent with the others; we have used all the data to establish a reliable calibration.

Figure 2. Differential number counts for stars for each plate in the north field. Note the abrupt change of slope at $r \approx 19.4$; this is due to contamination of galaxies being misclassified as stars. The large scatter from field to field is due to the fact that we are probing different regions of the Galaxy in each field.

Figure 3. Differential number counts for stars for each plate in the south field.

Figure 4. Star counts for each field in $17.5 \leq r \leq 18.5$ compared to the values predicted by the galactic model of Bahcall and Soneira. The triangles are for the south fields, and the filled squares for the north fields. The open squares are star counts in the north fields in the range $17.2 \leq r \leq 18.2$. The northern fields have systematically more stars, because their galactic longitudes are closer to the direction of the galactic center. (The latitudes of the two fields are similar.) The filled symbols fit the galactic model much better than the open squares plus filled triangles, thereby indicating that a magnitude error is unlikely.

Figure 5. The ratio of the correlation function $w(\theta)$ in the north field to that in the south field for small angles. Filled symbols are for the original catalogues, open symbols for the same south catalogue, but with the north catalogue selected 0.3^m brighter. The filled symbols are closer to unity than the open symbols.

Figure 6. Contour plot of the density of galaxies in the north field. The solid contour represents 400 galaxies per square degree (the global surface density of galaxies over both fields). The dotted contour is 20% under-dense and the dashed contour is 20% over-dense.

Figure 7. Same as Figure 6, but for the south field.

Figure 8. Differential number counts of galaxies are shown for the north and south fields. All the plates in each field have been averaged. The error bars due to counting statistics are smaller than the symbols, and so are not shown. The magnitude calibration is not reliable for objects with $r \leq 16.5$. Note the incompleteness near $r \approx 19.5$. The large discrepancy (40% at $r = 18.0$) is real; only 10% of the overdensity can be explained by differential stellar contamination of the sample.

Figure 9. Differential number counts of galaxies are shown for each individual plate in the north field. The legend indicates which fields are shown. Notice the open triangle, the large deviation from the average is due to the presence of the Corona Borealis supercluster. Field 387, denoted by the crosses, is not reliable (see Table 2.) but does not stand out markedly in this plot.

Figure 10. Same as Figure 9, but for the south field. The field to field scatter is somewhat larger than in the north.

Figure 11. The fraction of galaxies observed in the survey contained within a distance D in a model where the galaxies are distributed randomly in a cone of length $1000h^{-1}$ Mpc. A Schechter function with $M_* = -20.73$ was used.

Figure 1.

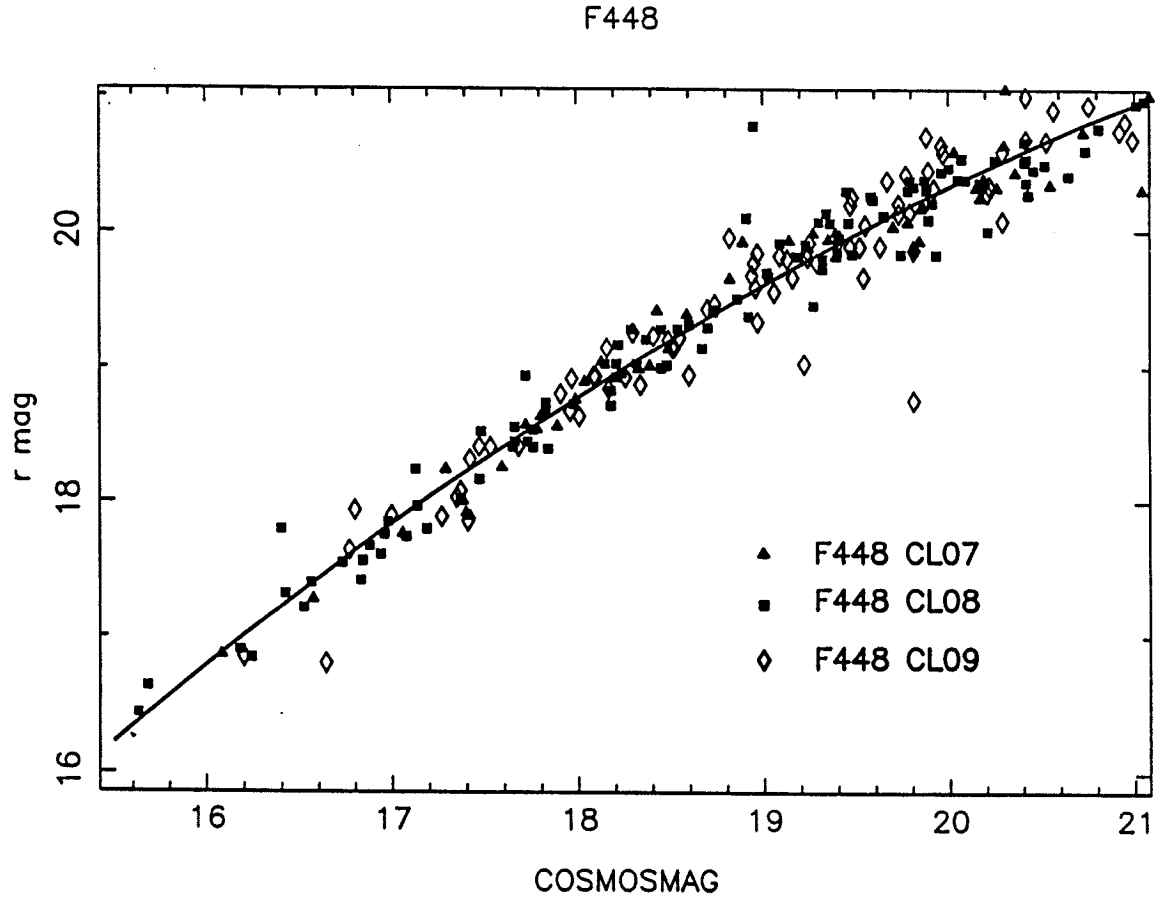

F448
r mag
16
18
20
16
17
18
19
20
21
COSMOSMAG
F448 CL07
F448 CL08
F448 CL09

Figure 2.

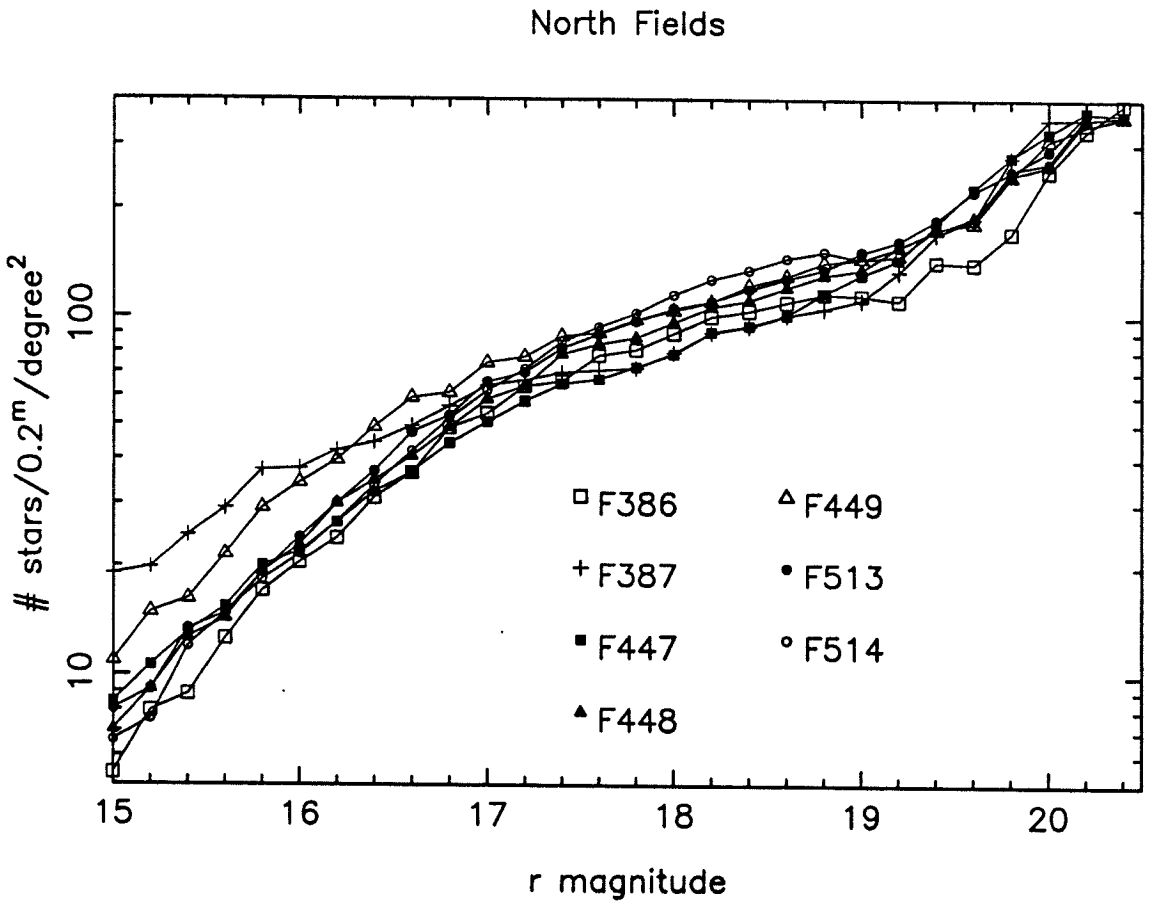

North Fields
stars/0.2^m/degree^2
10
100
15
16
17
18
19
20
r magnitude
F386
F387
F447
F448
F449
F513
F514

Figure 3.

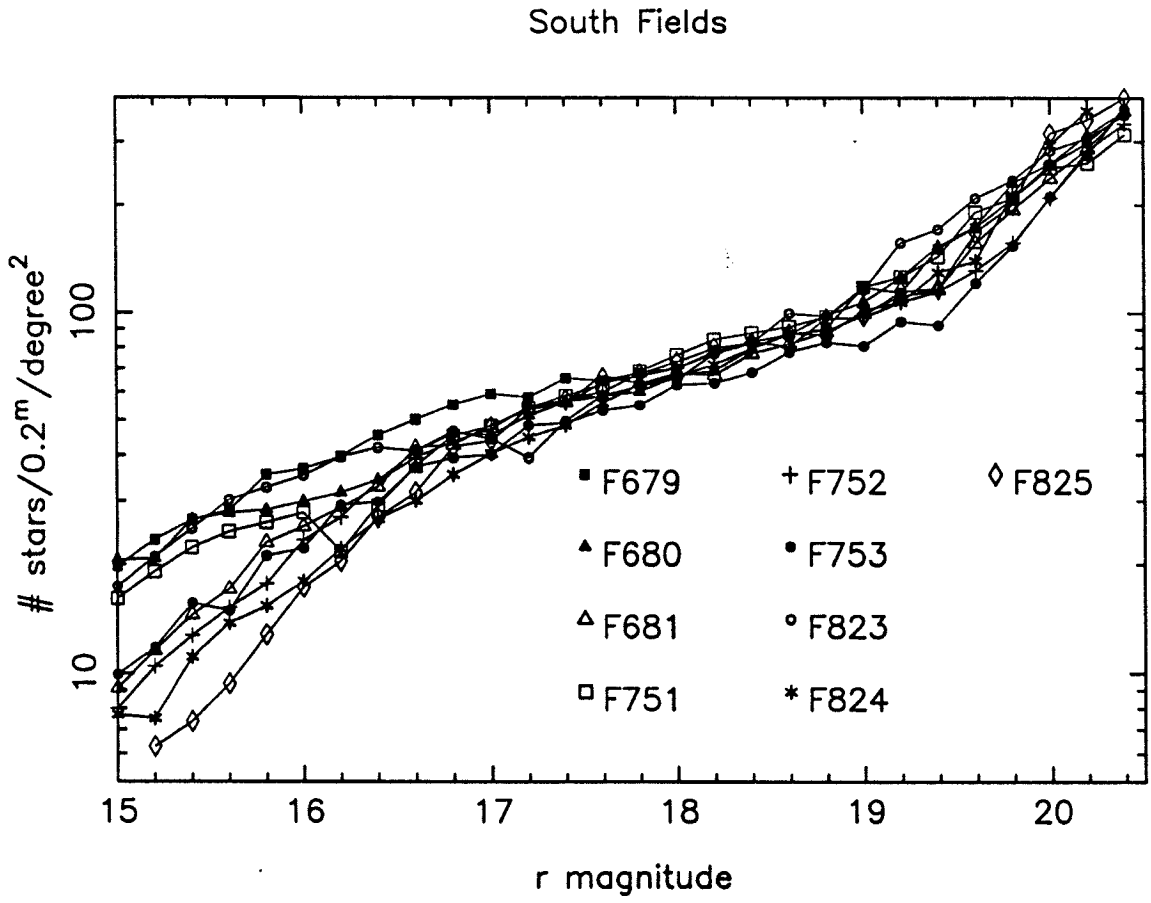

South Fields
stars/0.2^m/degree^2
10
100
15
16
17
18
19
20
r magnitude
F679
F680
F681
F751
F752
F753
F823
F824
F825

Figure 4.

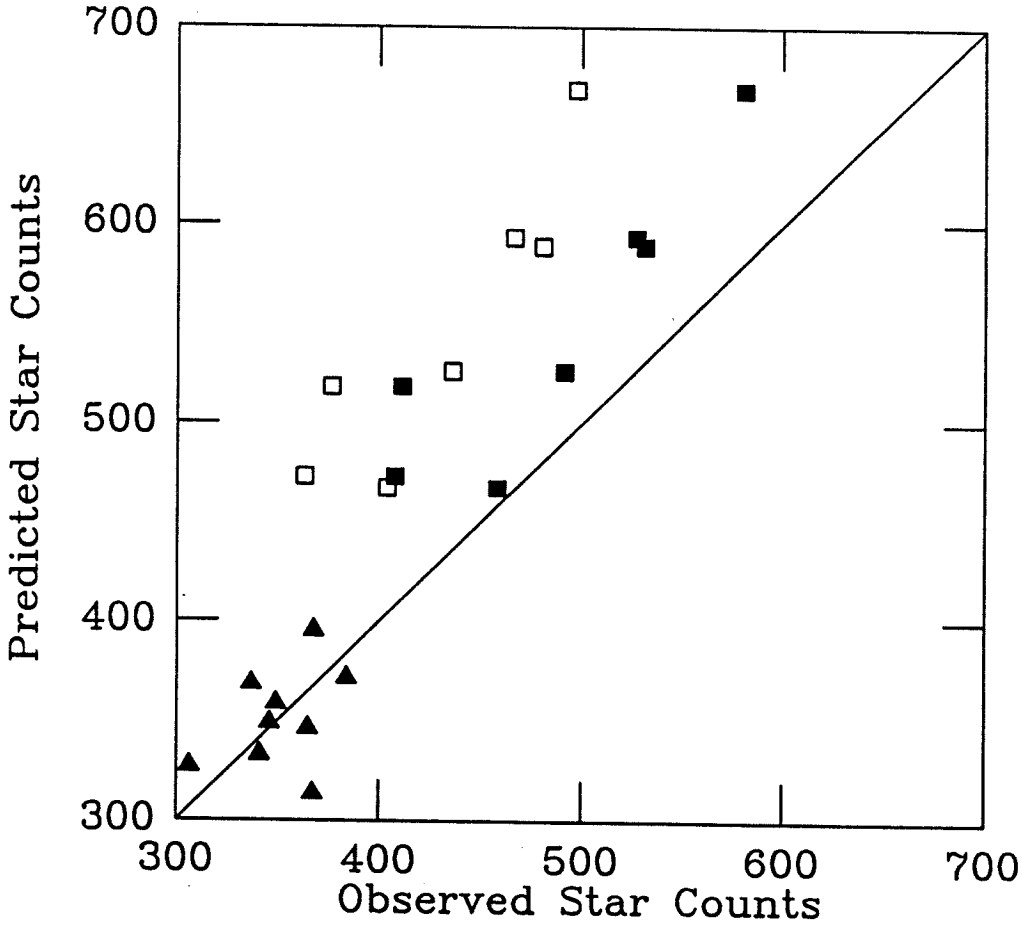

700
600
500
400
300
Predicted Star Counts
300
400
500
600
700
Observed Star Counts

Figure 5.

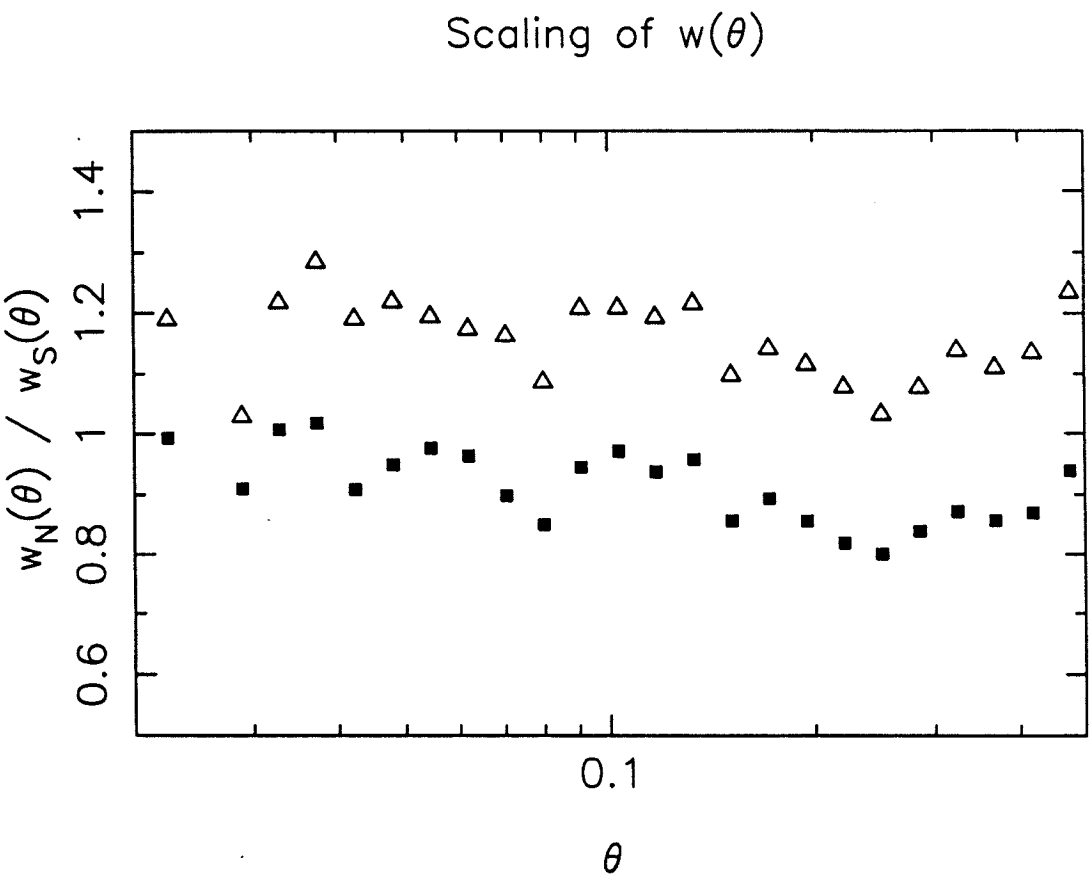


Figure 6.

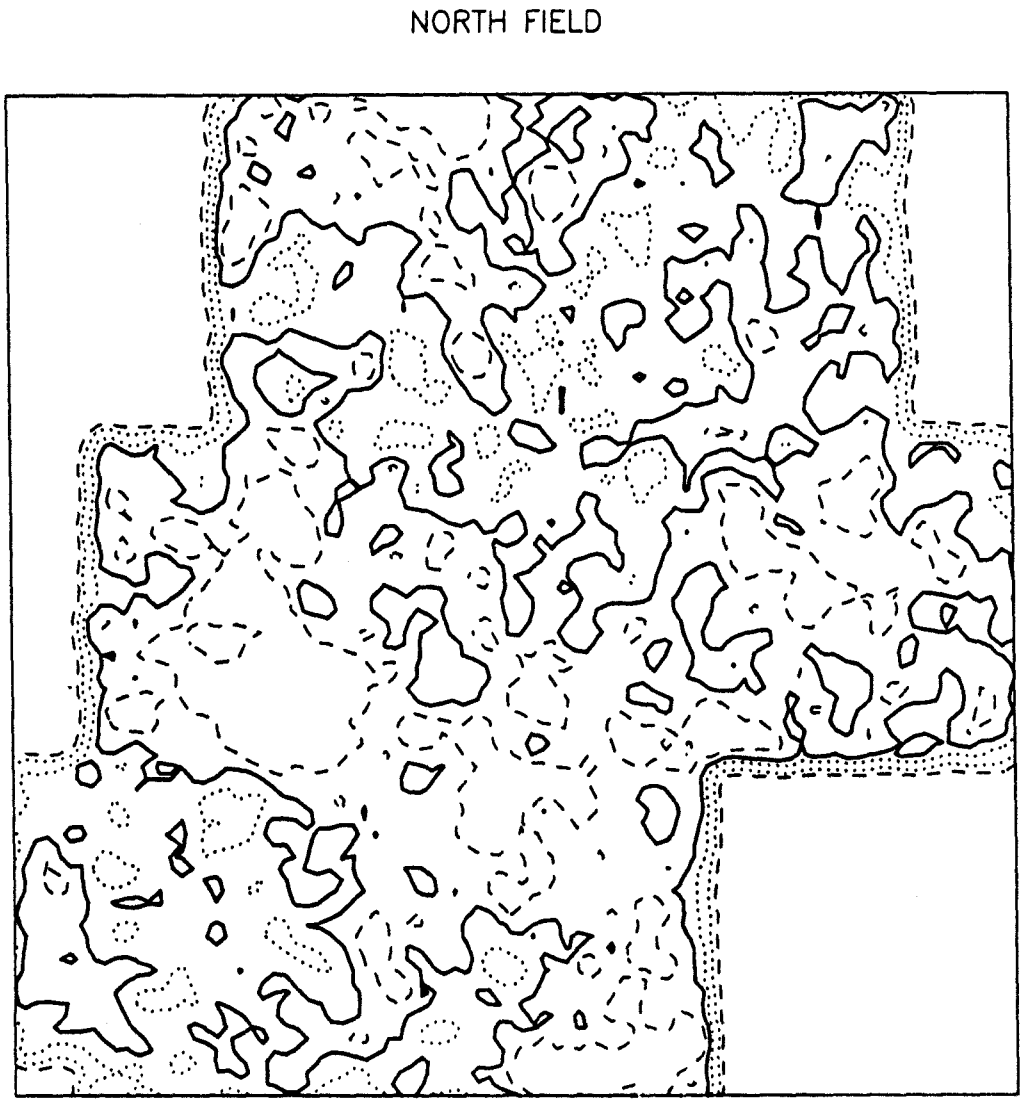


Figure 7.

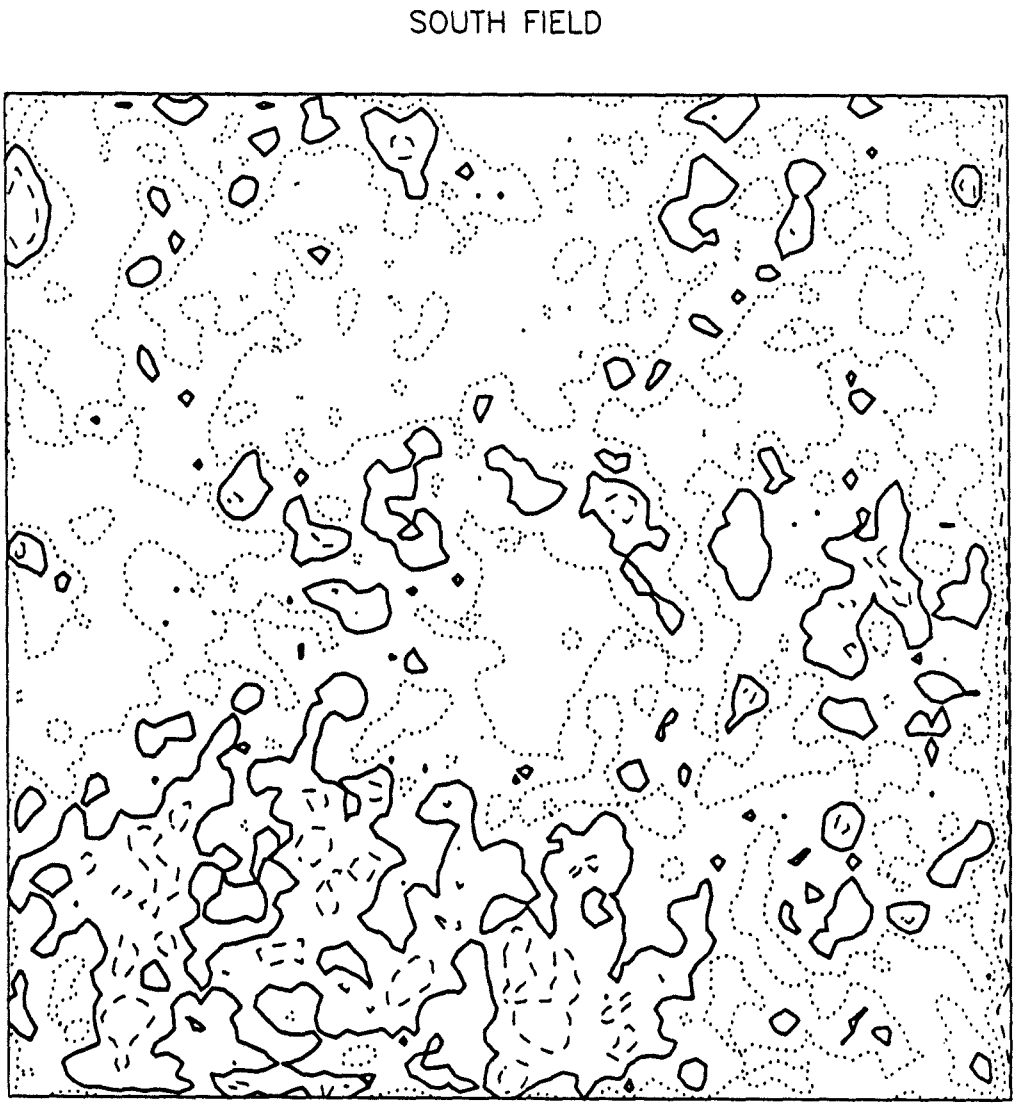


Figure 8.

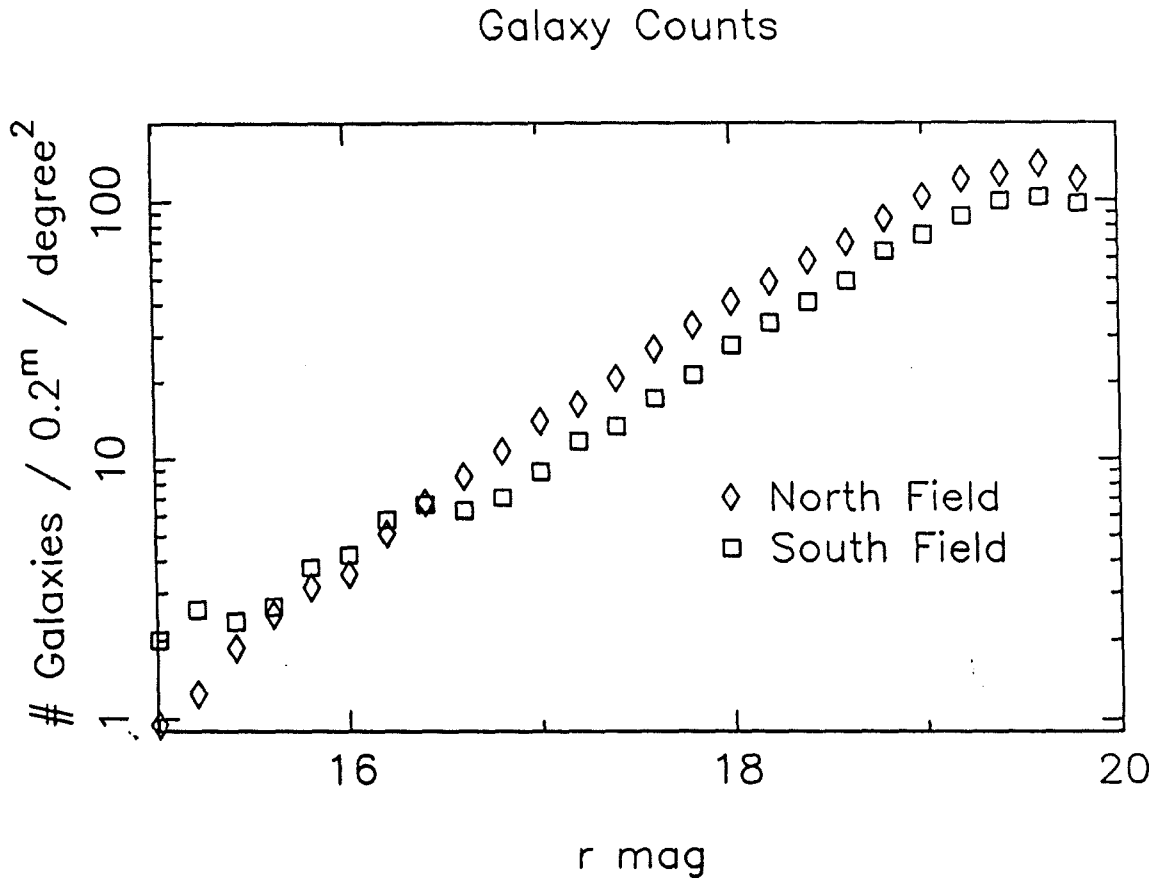

Galaxy Counts
Galaxies / 0.2^m / degree^2
1
10
100
16
18
20
r mag
North Field
South Field

Figure 9.

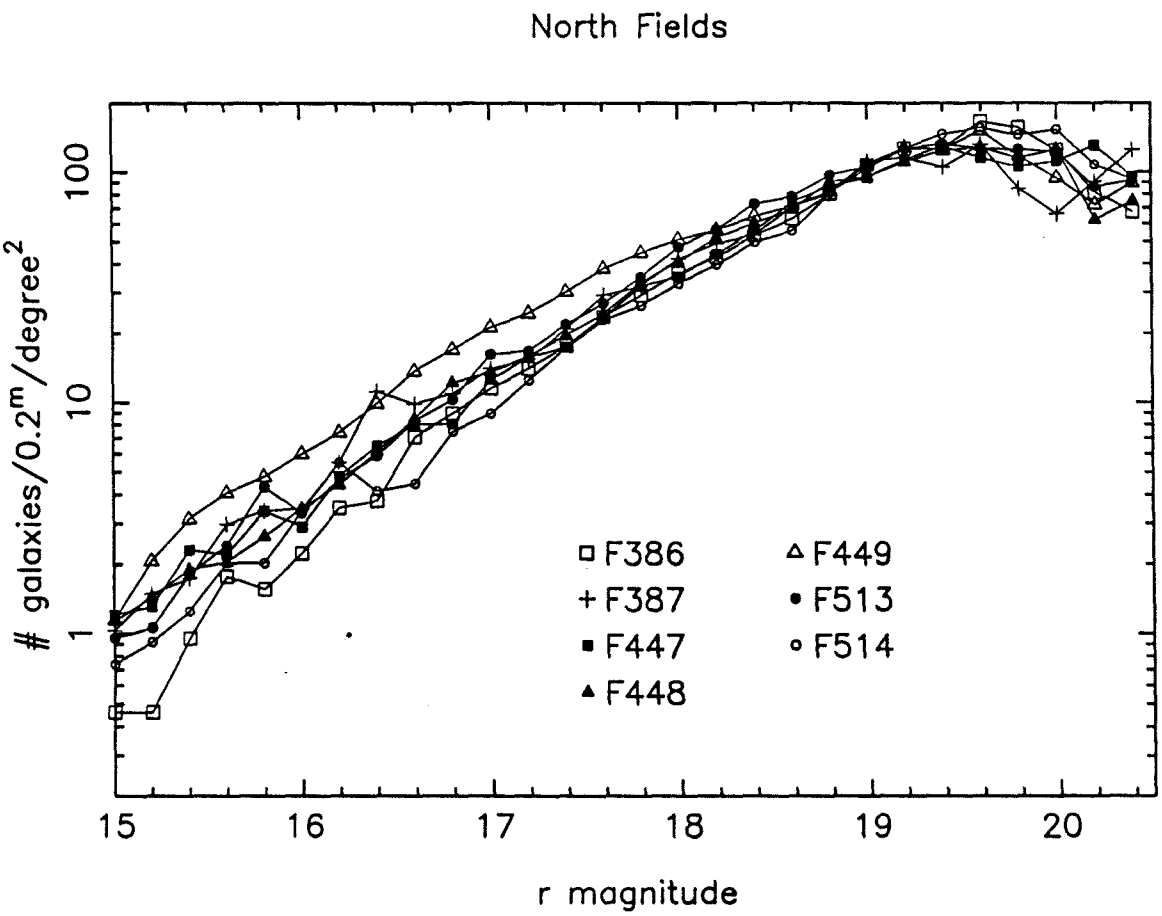

North Fields
galaxies/0.2^m/degree^2
1
10
100
15
16
17
18
19
20
r magnitude
F386
F387
F447
F448
F449
F513
F514

Figure 10.

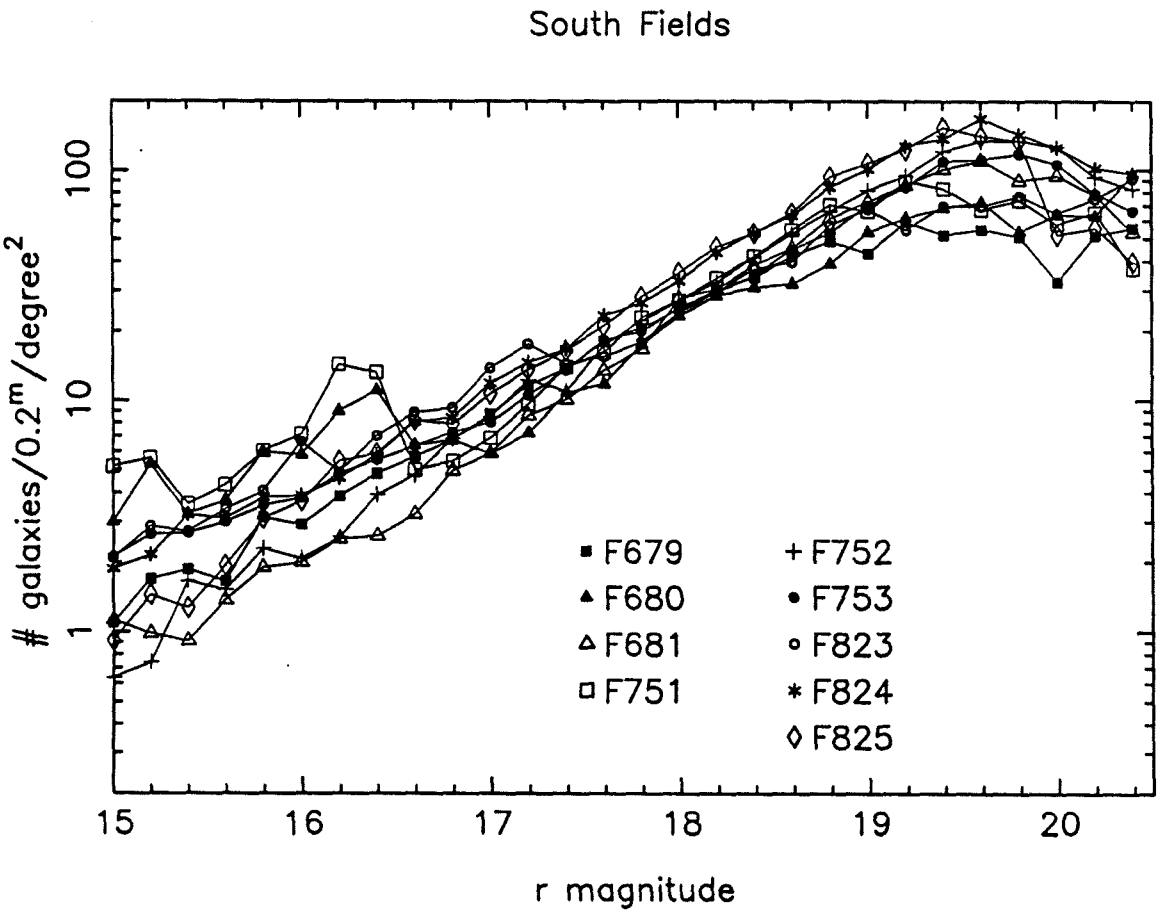
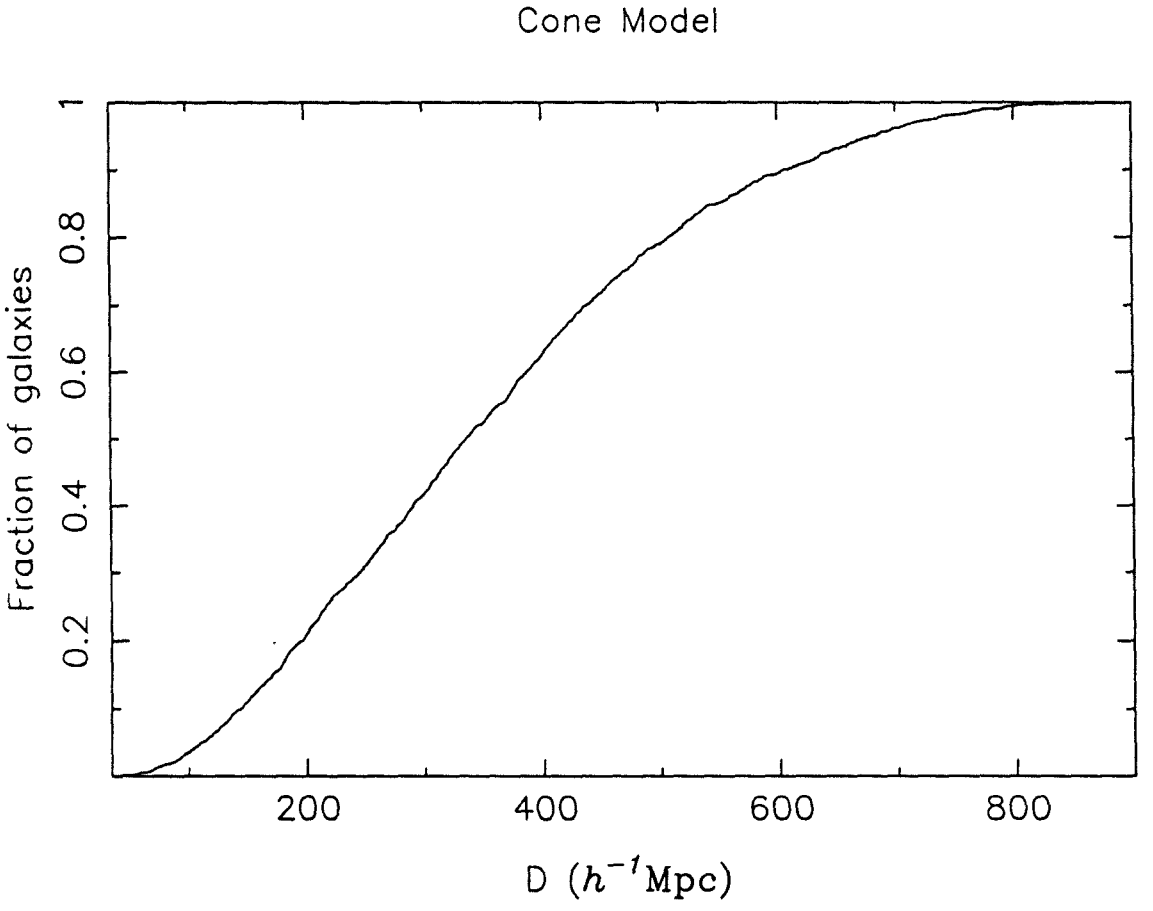

South Fields
galaxies/0.2^m/degree^2
1
10
100
15
16
17
18
19
20
r magnitude
F679
F680
F681
F751
F752
F753
F823
F824
F825

Figure 11.


Cone Model
Fraction of galaxies
0.2
0.4
0.6
0.8
1
200
400
600
800
D (h^{-1}Mpc)

4.7 References

Bahcall, J.N. and Soneira, R.M., 1980, *Astrophys. J. Suppl.*, **44**, 73.

Bahcall, N.A. and Soneira, R.M., 1984, *Astrophys. J.*, **277**, 27.

Broadhurst, T.J., Ellis, R.S., Koo, D.C. and Szalay, A.S., 1990, *Nature*, **343**, 726.

Burstein, D. and Heiles, C., 1982, *Astron. J.*, **87**, 1165.

Cowie, L. L., Gardner, J. P., Lilly, S. J., and McLean, I., 1990, *Astrophys. J. (Letters)*, **360**, L1.

Efstathiou, G., Ellis, R. S. and Peterson, B. A., 1988, *M.N.R.A.S.*, **232**, 431.

Geller, M. J. and Huchra, J. P., 1988, *Large-Scale Motions in the Universe: A Vatican Study Week*, Rubin, V. and Coyne, S. J., *ed.* Princeton University Press.

Geller, M. J., de Lapparent, V. and Kurtz, M. J., 1984, *Astrophys. J. (Letters)*, **287**, L55.

Gott, J. R., Melott, A. and Dickinson, M., 1986, *Astrophys. J.*, **306**, 311.

Groth, E.J. and Peebles, P.J.E., 1977, *Astrophys. J.*, **217**, 385.

Juszkiewicz, R., Krzysztof, G. and Silk, J., 1987, *Astrophys. J. (Letters)*, **323**, L1.

Kent, S. M., 1985, *Pub. Aston. Soc. Pac.*, **97**, 165.

Kirshner, R. P., Oemler, A., Schechter, P. L. and Shectman, S. A., 1981, *Astrophys. J. (Letters)*, **248**, L57.

de Lapparent, V., Geller, M. J., and Huchra, J. P., 1988, *Astrophys. J.*, **332**, 44.

Limber, D. N., 1954, *Astrophys. J.*, **119**, 655.

MacGillivray, H. T. and Stobie, R. S., 1984, *Vistas in Astronomy*, **27**,433.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989a, *M.N.R.A.S.*, **242**, 43p.

Meinhold, P. and Lubin, P., 1991, *in press.*

Peebles, P. J. E.,*The Large Scale Structure of the Universe*, Princeton University Press, 1980.

Picard, A., 1991*a*, *Astrophys. J. (Letters)*, **386**, L7—L10.

Picard, A., 1991*b*, thesis, in preparation.

Postman, M., Geller, M. J. and Huchra, J. P., 1986, *Astron. J.*, **91**, 1267.

Reid, N., 1990, *M.N.R.A.S.*, **247**, 70.

Schechter, P., 1976, *Astrophys. J.*, **203**, 297.

Schneider, D. P., Gunn, J. E. and Hoessel, J. G., 1983, *Astrophys. J.*, **264**, 337.

Seldner, M., Seibers, B., Groth, E. J. and Peebles, P. J. E., 1977, *Astron. J.*, **82**, 249.

Shane, C. D., and Wirtanen, C. A., 1967, *Pub. Lick Obs.*, **22**, Part 1.

Smoot *et al.* 1991, COBE preprint no. 90-07.

Thuan, T. X. and Gunn, J. E. G., 1976, *Pub. Aston. Soc. Pac.*, **88**, 543.

Tyson, A.J. and Seitzer, P., 1988, *Astrophys. J.*, **335**, 552.

Jarvis, J. F. and Tyson, A.J., 1979, *SPIE Proc. Instrumentation in Astronomy*, **3**.

Valdes, 1982, *Instrumentation in Astronomy* IV, SPIE **331**,465.

5. An Objective Catalogue of Clusters of Galaxies

5.1 Introduction

An objective catalogue of clusters of galaxies has long been known to be desirable. To quote from the paper of Abell, Corwin and Olowin (1989):

> "We hope that this will be the last such catalog prepared by visual scans of photographic plates, and we urge future investigators to compile cluster catalogs using high speed microphotometric scanning machines and objective selection criteria."

Previous attempts at such a catalogue include those of Shectman (1985) and Dodd and MacGillivray (1986). Unfortunately, the Shectman catalogue is based on the Lick counts, which suffer from variations in the depth of the plates. The catalogue of Dodd and MacGillivray, consisting of only one Schmidt plate, is too small to be of use for statistical analysis. Both these catalogues employ a clustering algorithm which is strictly an overdensity criterion. There is presently an effort to obtain an objective catalogue from the APM survey (Maddox *et al.* 1989 *a* and *b*), but no results have been published yet.

For several years now, clusters of galaxies have been used as a tool for measuring the large-scale structure of the universe. Perhaps one of the most interesting results from this method is that of Bahcall and Soneira (1983, hereafter BS83), who, in their seminal paper, analyzed a subsample of the Abell catalogue and concluded that the amplitude of the cluster-cluster spatial correlation function was 20 times as large as that of galaxies. Such power on large scales provides a formidable constraint for theories of formation of structures. At the present, no simple theory is capable of accommodating this result. The currently fashionable model of cold dark matter (CDM) with $\Omega = 1$ is unable to reproduce such power on large scales. Efstathiou *et al.* (1990) have recently suggested that a cosmological constant may provide a solution to this problem.

The BS83 result is based on an analysis of 104 Abell clusters with distance class D $\leq$ 4 clusters. It has been pointed out by Postman, Geller and Huchra

(1986) that the result of BS83 is sensitive to the presence of the Corona Borealis supercluster, in the sense that the removal of only six clusters from BS83's sample removes most of the correlation between 10 and $20h^{-1}$ Mpc. Thus they suspect that the BS83 sample is not large enough to be a fair sample. BS83 provides as additional evidence for their result that the D = 5 clusters, with redshifts estimated from the tenth brightest magnitude, support their result. Unfortunately, Postman, Geller and Huchra (1986) point out that the number of interlopers within an Abell radius at a redshift of 0.1 is 10, so that the estimated redshifts may be biased low, adding to the confusion.

There has recently been much discussion of the possible systematic errors in the Abell catalogue due to projection effects and how they affect the result of BS83. The main effect is that of a foreground cluster artificially enhancing the richness of a poor background cluster which would not have been included in the catalogue had it not been for the presence of the foreground cluster. Sutherland (1988) presents evidence for projection effects in the Abell catalogue by examining the redshift correlation function, and he concludes that $\xi(r) = 0$ for $r \geq 50h^{-1}$ Mpc.

Dekel *et al.* (1989) arrive at a similar conclusion by examining a "decontaminated" version of the BS83 sample. The decontamination method, however, is model dependent (these authors claim that the dependency on the exact model is weak) and depends on the galaxy-cluster correlation function. The model is one where there are clusters present in a uniform background (much as in our toy model in Chapter 3). Given the subtlety of the problem addressed, it is unclear what the effects are of using such an unrealistic model.

To confound matters further (!) West and Van den Bergh (1991), from a study of clusters containing a cD galaxy, have concluded that the original result of BS83 is probably correct. The rationale here is that cD clusters are more easy to identify, and thus less likely to suffer from the projection effects mentioned above. Unfortunately, cD clusters may not be typical; as it may well be that only very rich clusters evolve to produce a cD galaxy. It is know that there is a strong richness effect in the cluster-cluster correlation: richer clusters correlate more strongly (BS83). Thus the

strong $\xi(r)$ for the cD clusters may simply reflect the fact that rich clusters seem to have a higher correlation amplitude than poor ones.

This whole question will remain difficult to resolve until an adequate sample from X-ray selection criteria can be analyzed. In such a sample, the clusters are not subject to projection effects. Such a sample will be compiled from the ROSAT all-sky survey, and will hopefully be available soon.

In this chapter we will present a new catalogue of clusters of galaxies, defined using an objective algorithm. We hope that this catalogue, being constructed from a highly uniform sample of galaxies, will help in resolving some of the problems mentioned above. Our catalogue, being of rather limited angular extent, cannot of course be considered definitive. It is nevertheless large enough to serve as a base for statistical analysis. In particular, a comparison of the results from our catalogue to the Abell catalogue will help us evaluate how much weight should be given to previous results. Furthermore, a comparison between this catalogue and an X-ray selected catalogue should give us insight on how much trust can be put on catalogues of clusters derived from two-dimensional data, even when this data is of the highest possible quality.

Our chief aim in this chapter, then, will be to present in detail the algorithm used to construct our objective catalogue of clusters of galaxies, and to examine its statistical properties. This catalogue will also be used to assess the completeness of the Abell catalogue.

5.2 Cluster Finding Algorithm

5.2.1 Abell's Method

Of the existing catalogues of clusters of galaxies, that of Abell (1958) is the one which has been used most extensively for statistical analyses. It is therefore worthwhile to review the method used by Abell for the construction of his catalogue. An understanding of his method will enable us to adapt part of his procedure in a rigorous (*i.e.* , objective) way, and to understand the methodological differences in the procedures which lead to disagreement between the two catalogues.

Abell visually inspected all plates, and flagged interesting objects. It is not specified if this was done in some systematic way (by first drawing a grid of small boxes on the plate, for example). Then each cluster was tested against the following criteria:

1. *Richness Criterion.* Each cluster must contain at least 50 members that are not more than 2 magnitudes fainter than the *third* brightest member galaxy.
2. *Compactness Criterion.* A cluster must be sufficiently compact that these 50 galaxies are within a (previously agreed upon) radial distance†.
3. *Distance Criterion.* A cluster must be sufficiently distant that it does not extend over more than one (or at most two) plates.
4. *Galactic Latitude Criterion.* Abell specifies which region of the sky he feels can be surveyed in a complete manner, without danger of confusion from stars in the galactic plane, or incompleteness due to galactic obscuration.

He estimated the magnitudes visually by comparing with a film copy of several standard galaxies. The redshift of the cluster was estimated using a magnitude–redshift relation for the tenth brightest galaxy. This yielded the appropriate angular radius over which he included galaxies. From the number of galaxies in this circle and magnitude range, Abell subtracted what he estimated to be the contamination

† This is now know as an *Abell Radius.* Abell was using a Hubble constant of 180 km/s/Mpc. In modern terminology, an Abell radius is $r_A = 1.5h^{-1}$ Mpc.

from the field, which he obtained by looking at regions on the same plate judged to be devoid of clusters. This correction can be as large as 30%.

Finally, by comparing lists obtained from repeat plates, he estimated that no more that 2% of the clusters identified on a preliminary inspection were missed on final inspection.

This procedure is, at least in theory, reasonably robust. In particular, the use of the "magnitude strip" is an attempt to ensure that the measured richness of the cluster is independent of its distance. The *relative* measurement of the magnitudes is reasonably good (as indicated by the repeat measurements), and the exact form of the distance-redshift relation used is not very important. Variations in the quality of the plate material, however, render the *absolute* measurement of the magnitudes highly suspect. This effect can easily be seen to lead to a lack of uniformity in the depth surveyed across the sky. This is obviously an area where our catalogue, being constructed from photometrically very accurate and uniform data, will be superior. A second, equally important pitfall of this method, is of course the lack of objectivity in the procedure for deciding if an overdensity of galaxies on the sky corresponds to one or more clusters. In particular, consider the case where two or more clusters are close together on the sky. What should be the criterion for deciding if the observed overdensity of galaxies is in fact one or two clusters? If there is a small concentration of faint galaxies superposed on a larger, looser concentration of brighter ones, we may reasonably conclude that we are detecting two clusters, at very different redshifts. For objects that are at comparable distances, however, the case is not so clear cut. Examples of clusters with sub-clustering and multiple nuclei abound (Rhee 1989), as well as examples where detailed redshifts reveal two (or more) distinct clumps in redshift space (Struble and Rood, 1987) in an Abell cluster (presumably indicating two or more physical clusters seen in projection).

Some of these problems can be addressed using an objective algorithm; others are intrinsic to any method which does not have access to full three-dimensional information.

5.2.2 The Galaxy Density Field

Our cluster finding algorithm is based in part on that described by McGill and Couchman (1990), hereafter MC90, which is in turn based on Abell's method. In their paper, they describe a procedure which models that of Abell, and then proceed to suggest methods of improving the scheme. MC90 were especially interested in calculating the "signal to noise" of detection of a cluster. The key to their paper is in deriving an expression for the "noise", which involves an integral over the two point correlation function of galaxies w(θ) (derived in Chapter 3). Unfortunately, since the clusters themselves contribute to w(θ), it is unclear how much of the "noise" being subtracted is in fact a real signal.

The first step in our (and MC90's) procedure is to produce a smoothed density field starting from the discrete galaxy catalogue. The density field of a galaxy catalogue of N objects can be represented as a summation of delta functions, thus,

$$\rho(\Omega) = \sum_{i=1}^{N} \delta(\Omega - \Omega_i) \tag{1}$$

where Ω_i are the angular coordinates of the i^{th} galaxy.

This density field is then smoothed as follows:

$$\rho_s(\Omega) = \int \rho(\Omega')S(\Omega, \Omega')d\Omega' \tag{2}$$

Equations (1) and (2) are taken straight from the appendix of MC90, and we have kept their notation to minimize confusion.

We have implemented these formulas on our catalogues using a truncated gaussian smoothing filter,

$$S(r) = \begin{cases} Ce^{(r/R_{\text{smooth}})^2} & \text{if} \quad r \leq 2R_{\text{smooth}} \\ 0 & \text{if} \quad r > 2R_{\text{smooth}} \end{cases} \tag{3}$$

where r is the angular distance between the two points being considered. The filter is truncated for computational reasons. At $r = 2R_{\text{smooth}}$, the filter has fallen to $e^{-4} \approx 1.8\%$ of its peak value.

This smooth density field is then sampled on a rectangular grid of pixel size Δx, and stored in the standard FIGARO image format, enabling us to use the existing battery of programs to display and analyze the image. We have to decide on a pixel size Δx on which to sample our density field $\rho(\Omega)$. Our criterion was that Δx should be small enough that the density field be oversampled, *i.e.* , that the chosen pixel size should be much smaller than any smoothing scale to be considered. On the other hand, we do not wish the resultant image to have enormous dimensions, as this would render the subsequent analysis difficult: hence the pixel size must be as large as possible. Also, it is convenient to keep this scale constant for all smoothing radii, so as to simplify the subsequent steps in the analysis. The surface densities of galaxies for our north and south galaxy catalogues ($16.5 \leq m_r \leq 19.0$) are 482 and 334 galaxies per square degree respectively, which correspond to typical interparticle distances of 185″ and 224″. At the limiting magnitude of our catalogue, an L_* galaxy is located at a distance of $880h^{-1}$ Mpc (see § 4.5). The angle subtended by a $1h^{-1}$ Mpc physical feature (such as a cluster) at this distance is 235″. Since an Abell radius is $1.5h^{-1}$ Mpc, this is the smallest angular distance we would want to smooth on. We therefore, somewhat arbitrarily, chose to sample our density field on a grid of pixels separated by 100″. This grid size yields images which are 512×525 pixels for the north field, and 526×537 for the south.

In practice, we implemented Equation (2) as follows. The problem is to evaluate the density field `rho(i,j)`[†], summing the contributions from each galaxy. It is computationally too expensive to apply Equations (2) and (3) directly, because the exponential needs to be calculated on the order of $(R_{\rm smooth}/\Delta x)^2 \times$ Npts times. Instead, we define a square matrix `filter(i,j)` in which we load the values of the smoothing filter, where the distance of the point `i,j` is measured from the origin, *i.e.* , $r^2 = \mathtt{i}^2 + \mathtt{j}^2$. The indices `i,j` are free to run from $-2R_{\rm smooth}/100$ to $+2R_{\rm smooth}/100$, (where $R_{\rm smooth}$ is measured in arc-seconds), thus implementing Equation (3). The normalization constant C in Equation (3) is automatically chosen by the program so that the weight of the filter (defined as $\sum_i \sum_j$ `filter(i,j)`) is

† For clarity, we denote the variables used in our FORTRAN programs with the `typewriter` font.

equal to 100[‡]. Thus, one galaxy will contribute 100 units of "flux" to our smoothed density field.

We then look through each galaxy in the catalogue, compute its coordinates `i,j`, and add the (small) matrix `filter(i,j)` to the relevant submatrix of `rho`, as follows:

$$\texttt{rho(i,j)} = \sum_{\substack{k=\max(1,\\ \texttt{i-NFILTMAX})}}^{\min\left(\substack{\texttt{NXMAX},\\ \texttt{i+NFILTMAX}}\right)} \sum_{\substack{l=\max(1,\\ \texttt{j-NFILTMAX})}}^{\min\left(\substack{\texttt{NYMAX},\\ \texttt{j+NFILTMAX}}\right)} \texttt{filter(k,l)} \qquad (4)$$

where `NXMAX,NYMAX` are the dimensions of `rho(i,j)`, `filter(i,j)` is declared as `filter(-NFILTMAX:NFILTMAX,-NFILTMAX:NFILTMAX)`, and `NFILTMAX` $= 2R_{\rm smooth}/100$.

The computational savings of this approach (compared to the brute force approach of calculating the value of S at each relevant point) are enormous; the execution time drops from over an hour of CPU on the CONVEX to 30 seconds. There is of course a penalty; in this case a loss of resolution. The filter behaves as if each galaxy is in the middle of the cell, which is usually not the case. However the error introduced is small, because $R_{\rm smooth} \gg \Delta x$.

An example of the results obtained from this program is shown on Figures 1, 2 and 3. The latter two Figures are for smoothing lengths of 300″ and 600″. The area shown is 5000″ on a side, and is taken from Field 681. Figure 1 shows the position of every galaxy in the test field, whereas Figures 2 and 3 show a grey scale map of the surface $\rho(\Omega)$. Note the presence of the two clusters of galaxies, Abell 68 and Abell 73. The clusters are more apparent on Figure 3 than on Figure 2 because the smoothing scale is better matched to the clusters' redshifts. Although these can easily be recognized when pointed out, the density contrast in the dot plot is low; the clusters are unmistakable on the smoothed density map, however.

Plate 1 (the frontispiece of this volume) is also an example of this smoothing method. Here, only galaxies in the magnitude range $17.5 \leq r \leq 18.5$ are shown,

‡ We could have chosen the filter weight to be unity, but the output files were stored in FIGARO format, and it is simpler to handle images whose dynamic ranges are reasonably large.

and the smoothing length is 420″. The irregularly shaped frame (to the right) is the north field. In this orientation, north is up and east to the left. The angular extent of each image is approximately 15°. The contrast is such that individual galaxies are invisible; the features seen on this plate are all groups of galaxies, with the dimmest visible features being composed of 3 or 4 galaxies.

5.2.3 Detection of Peaks

The next phase consists of locating the sites of the candidate clusters. In essence, we wish to locate all of the local maxima of the surface $\rho(\Omega)$. We first experimented with some well known detection schemes, such as the *detect* algorithm from FOCAS (see § 2.3.4). To our surprise, these did not work very well, *i.e.* , they miss several peaks which are readily identified by visual inspection. The reason for this is the highly non-gaussian nature of the "background" in the image. The FOCAS algorithm (and others like it) assume that the background is essentially constant, with some gaussian fluctuations. In our frames, however, the strong two point correlation function of galaxies causes the background to vary substantially, and to have a highly non-gaussian appearance. Figure 4 shows the distribution of pixel values for the north and south images with $R_{\rm smooth}= 420''$; notice the strong tail for large data values. The north field is the curve with the lower peak and higher tail.

We therefore coded our own detection algorithm. It is rather simplistic, but makes no assumptions about the background and so is well suited to our problem. First an initial threshold is chosen. In practice, this was done by examining a histogram of the pixel values on the image being processed (as shown on Figure 4). The threshold is then set to be that value V where the number of pixels with value V is 80% that of the number of pixels with the modal value (and $V >$ mode). This threshold is sufficiently small that even the smallest of peaks likely to be clusters are detected. Every pixel which satisfies this criterion is then flagged, and the pixels that are neighbours to each other (using 8-point connectivity) are put in a list defining an "object". For each such object, the x, y coordinates (intensity weighted centroid) in the image reference frame, area (number of pixels) and threshold level are written to a file. The threshold is then increased by a fixed amount (in our case,

5%) and the procedure is repeated until the threshold is so large that no pixels are flagged.

Obviously, in this procedure, an isolated, bright object is found on several passes. Where two objects are close together, however, they are first detected as only one object when the threshold is below the saddle point (on the $\rho(\Omega)$ surface) joining the two objects, but eventually get separated out. This algorithm, then, works very much like that of the COSMOS deblending software (Chapter 2); however, its implementation is more rudimentary.

It then remains to go through the output list and remove all redundant objects. This is done by starting at the beginning of the list, and checking against every object below the current object in the list to see if the second object is sufficiently close to be the same, and then keeping only the object found *at the highest threshold.* Two objects are deemed to be "sufficiently close together" if their distance does not exceed $d = \sqrt{\text{area}/2}$. We do this because merged images are first detected as elongated blobs, with the centroid often near the saddle point between the two objects.

The final list thus consists of all the positions representing a local maximum in the density field which have a sufficiently bright core. For each of these positions, we sum the flux in a circular aperture whose radius is equal to the smoothing radius which was used to generated the smoothed density field. We denote this quantity $\mathcal{L}$†. For this we use a modified version of the FOTO program, which is part of FIGARO. The modification is trivial and one of simplification; we simply do not subtract the sky annulus and do not take -2.5 log of the computed luminosity. This program is especially useful for its ability to properly calculate the contribution of the pixels which are only partially inside the aperture.

5.2.4 Generating a List of Candidate Clusters

The galaxy catalogues are smoothed over a range of smoothing radii $R_{\rm smooth}$.

† We use $\mathcal{L}$ as a mnemonic for luminosity, although it should be kept in mind that this quantity represents a number of galaxies, and has nothing to do with how bright those galaxies are. The author apologizes for this confusing terminology.

The purpose of this is to enhance the signal from features of a given (constant) physical size as it is moved at more and more distant redshifts. Here, of course, the physical size we have in mind is that of a typical cluster of galaxies. The exact values chosen are not too important, since we use a range of values. In fact, we will later demand that a cluster be found in several smoothed catalogues to be considered real. The values of $R_{\rm smooth}$ used (given in arc-seconds) are shown in Table 1 below. They correspond to the angular size subtended by a $1h^{-1}$ Mpc bar at the redshifts shown, in a standard Friedmann cosmology with $\Omega = 1$ and $q_0 = \frac{1}{2}$, and thus are slightly smaller than an Abell radius. A nearby cluster will then be detected only when using a large smoothing radius, whereas a distant cluster is detected with a small smoothing radius.

Table 1 Smoothing Radii

$\mathcal{L}_{\#}$	z	$R_{\rm smooth}('')$
1	0.16	558
2	0.18	510
3	0.20	474
4	0.22	444
5	0.24	420
6	0.26	396
7	0.28	378
8	0.30	366
9	0.32	348
10	0.34	336

The smoothing procedure described in the previous sections is followed for each of these values of $R_{\rm smooth}$. We then merge the lists together, and produce a table of "clusters" and their measured photometry. A curve of $\mathcal{L}(R)$ *vs.* $R_{\rm smooth}$ is hereafter called the cluster's *profile.* Note that the profile is an integrated quantity, *i.e.* , part of the data used to determine $\mathcal{L}(R_1)$ goes into the determination of $\mathcal{L}(R_2)$, hence it is not surprising that the profiles look very smooth.

As in § 5.2.3, this merging is done by comparing the position for all possible pairs of objects, and deciding if they are close enough to in fact be the same object. Here, however, the criterion is that if the distance between the two centroids is

smaller than the maximum of the smoothing radii for the two objects, then they are the same object. The automated decision on whether a cluster is single or double is thus made by these two merging procedures. Note that, intrinsic to this algorithm, two clusters cannot be located within $R_{\rm smooth}(10) = 336''$ of each other.

5.2.5 Analysis of the Cluster Profiles

The profiles are, in essence, the final measured parameters for the clusters. They represent the object's overdensity, measured on several different angular scales. By selecting objects satisfying a given criterion, we can generate lists of objects, namely, our objectively selected sample of clusters of galaxies.

To aid in the interpretation of these profiles, we obtained the typical profile observed when the smoothing procedure is applied to a sample of randomly distributed objects. We constructed random catalogues with the same geometry and number of objects as the galaxy catalogues. The smoothing procedure was done as with the real data, and aperture photometry was performed using the coordinates of the Abell clusters as centers (though any random set of points would have done). For each smoothing length, we noted the average and standard deviation of the values of $\mathcal{L}$. These are shown on Figure 5. The profiles follow a perfect power law $\mathcal{L} \propto {R_{\rm smooth}}^2$. This should not come as a surprise, since, by construction, the smoothing filter always has the same weight, and there are always several objects present inside a smoothing radius (since $R_{\rm smooth} \gg$ mean interparticle spacing). Therefore the luminosity $\mathcal{L}$ simply grows as the area of the aperture, which is proportional to the number of galaxies included in the aperture. The different amplitudes for the power laws for the north and south are simply due to the difference in the mean density of the random catalogues, which, by construction, have the same surface density as the galaxy catalogues.

5.3 Results

5.3.1 The Multiplicity Function

We shall use the random profiles discussed in § 5.2.5, and introduce a parameter f, which represents the overdensity of a cluster compared to the typical profile found on a random catalogue. Thus, we say that a cluster is detected *at overdensity* f if each of its measured luminosities $\mathcal{L}_i^{\rm cluster} \geq f\,\mathcal{L}_i^{\rm random}$, for each value of i. If a cluster is not present on level i (*i.e.* , at the value of $R_{\rm smooth}$ corresponding to $\mathcal{L}_{\#} = i$ in Table 1), the test is not applied. However, a cluster must be found on *at least* three levels to be considered real.

Figure 6 shows the multiplicity function, which we define as the number of objects detected at overdensity f or greater as a function of f. The solid and dot-dashed lines are the result for the north and south fields respectively (the meaning of the symbols will be explained in § 5.3.3). The functions agree reasonably well for large overdensities ($f > 2.0$), but diverge for smaller values of f. It should be kept in mind that each field is referenced against the random profiles generated with the appropriate density, so that a cluster with $f = 2$ has a greater number of galaxies detected in it in the north field than in the south. This is in effect defining a cluster as a fixed overdensity over the *local* density of galaxies. Note that the number of clusters was not corrected for the different areas surveyed in the north and south. The south field's area is 29% larger. Such a correction would not show up as a large effect on Figure 6, and would go in the sense of driving the objective cluster's multiplicity functions closer together, and that of the Abell clusters further apart.

The other approach is to reference *both* fields to the same profile, say the south random profile. In this case, a cluster with $f = 2$ has the same number of galaxies in it in each field. This is shown in Figure 7. Unless stated explicitly, however, when we write f, we mean the one referenced to each field's own random profile.

The multiplicity function is seen to be *extremely steep*. For values of $f > 1.5$, power law fits of the form $N(> f) \propto f^{-\nu}$ have indices $\nu = 6.3$ for the north field, and $\nu = 7.0$ for the south. This means that the error in the number of clusters

detected goes as $\frac{\delta N}{N} = -\nu\frac{\delta f}{f}$. Abell claims an accuracy of 17% in the determination of the number of galaxies belonging to a cluster. This implies a 120% variation in the number of clusters detected.

5.3.2 A Catalogue of Objectively Selected Clusters of Galaxies

The procedure followed in § 5.2 yielded a total of 610 positions in the north field and 820 in the south which were detected at least at three levels. For each of these, values of f were calculated, and these are the catalogues which were used to determine the cluster multiplicity function.

We now present part of this data explicitly in Tables 3 and 4. Only those clusters with a value of f above 1.5 are given, due to considerations of space (and the fact that most clusters with $f < 1.5$ are probably too poor to be of interest to the observer). The full list is available from the author in machine readable format upon request.

The clusters have been sorted in increasing right ascension, and we have tabulated the positions in 1950 coordinates. These were obtained by transforming the centroids (in the image frame coordinates) back to the coordinate system of COSMOS, then using the coordinate transformation stored in the COSMOS header block to obtain the right ascensions and declinations. We expect the position to be good to better that 1 pixel (100″). We also give the values of f for each cluster, as well as its computed "galaxy flux" (roughly in units of 100 galaxies) for each smoothing radius. Thus, $\mathcal{L}(1)$ corresponds to the first value of $R_{\rm smooth}$ listed in Table 1. Where no entry is given, the cluster was not detected at that smoothing radius.

5.3.3 Assessment of the Abell Catalogue

We now turn our attention to the completeness of the Abell catalogue. Because our cluster finding procedure is not simply an automatic version of Abell's, we cannot measure Abell's criteria on our clusters†. We can, however, measure the profiles of the Abell clusters, and compare them with the overall population of such

† In fact, only Abell could measure Abell's criteria on our clusters!

objects found in the catalogue. The correspondence (or lack thereof) is bound to be instructive.

We have measured the profiles for each Abell cluster located on the area surveyed. The results are presented in Table 5 for the north field, and in Table 6 for the south field. The columns successively indicate the Abell cluster number, whether or not this cluster is included in Abell's "statistical sample", the distance and richness classes of the cluster, and our measured value of f, determined relative to the uniform profiles shown on Figure 5. Also tabulated are the 10 profiles $\mathcal{L}(1)\ldots\mathcal{L}(10)$. There are 54 Abell clusters located on the regions surveyed in the north, and 31 in the south.

We first investigate the relationship between the Abell distance class (D) and f. Figures 8 and 9 show the profiles for Abell clusters in the north and south fields. The squares (and dashed line) indicate the average profile for all clusters with $\mathrm{D} = 6, \mathrm{R} \geq 1$. The triangles (and dot-dashed line) indicate clusters with $\mathrm{D} = 5, \mathrm{R} \geq 1$, and finally, the crosses (and dotted line) indicate clusters with $\mathrm{D} = 4, \mathrm{R} \geq 1$. The random profiles are shown as diamonds and solid line. Note the difference in slope between the profiles of the clusters and that of the random profile; this is a manifestation of the cluster-galaxy correlation function.

In the north, there is a good correlation between the distance class and the location of the profile, in the sense that the more nearby clusters (smaller D) contain more galaxies in a fixed aperture. Because we are looking at a *fixed* magnitude interval, the galaxies we sample are intrinsically fainter for nearby clusters. Since there are many more faint galaxies than bright ones, we expect the profile to move up in Figure 8, as is observed. In the south field, the correlation breaks down for distance classes 4 and 5, but there are only two clusters with $\mathrm{D} = 4$, so not too much should be made of this discrepancy.

Another point to note is that, for a given distance class (say $\mathrm{D} = 5$), the overall $\mathcal{L}$ values are much higher in the north than in the south. If Abell was truly detecting and classifying the same objects all over the sky, this would not happen. Even if we subtract off the random profiles (to mimic Abell's background correction) as in

Figure 10, there is still an appreciable difference in the number of galaxies which make up clusters *which are assigned the same properties.* We conclude that Abell was unable to accurately follow the procedure he had set out for himself; clusters with a given richness and distance class are comprised of a significantly different number of galaxies, even after background subtraction. Of course, the difference in background density between our two fields is truly outstanding, and may constitute a worst-case scenario. Also, we do not know what the quality of the plates was like; a substantial difference in the limiting magnitude between the north and south fields, for example, could have caused him to err in this manner.

We next examine the correlation between Abell's richness class (R) and f. Figure 11 shows the values of f for all Abell clusters in the north field. Clusters have been segregated into their respective distance classes, because we cannot compare values of f across distance classes. The size of the symbol is proportional to the richness class; thus there are 4 size of symbols, from very small (R = 0) to very large (R = 3). The spread of values of f within a given distance and richness class (*i.e.* , for clusters with the same parameters) is quite large, obliterating any possible trend.

We now refer back to Figure 6, which shows the multiplicity function of the objectively detected clusters in the north and south as the solid and dot-dashed lines, and that of the Abell clusters in the north and south as the squares and triangles. Two things become immediately apparent: first, that the multiplicity functions for the Abell clusters are very different in each field, and second, that both fields are severely incomplete, in the sense that many more objective clusters are detected at a given value of f than are found among the Abell catalogue.

The first point may be a reflection that objects with the same physical properties are not being classified as being identical. If the background correction had been done properly, for example, the squares on Figure 6 would move towards the left, possibly bringing the two fields into agreement. Also, the possibility of large scale structure cannot be excluded, *i.e.* , there really may *be* more clusters in the north field than in the south field. Note also how using the local uniform profile as

reference (in Figure 7) drives the multiplicity function for the Abell clusters even further apart.

The second point is more serious. The figure shows that, even if we consider the most obvious, dense clusters (say $f > 2$), only a small fraction of those are recognized by Abell. This point is so important that we summarize the data of Figure 6 in the following table:

Table 2 Multiplicity Function of Clusters

f	N_A^n	N_{obj}^n	N_A^s	N_{obj}^s
3.00	2	3	1	2
2.75	3	4	2	4
2.50	4	5	2	6
2.25	6	8	4	13
2.00	18	32	9	34
1.75	28	58	14	89
1.50	36	119	22	212

where the subscripts 'A' and 'obj' denote the Abell and objective clusters, and the superscripts 'n' and 's' denote the north and south fields. The completeness is worse in the south, where, for $f = 2$, only 23% of the clusters are common to both catalogues!

Plates 2 and 3 show the location of the Abell clusters, denoted as circles, superposed on the smoothed density field, using the same parameters as for Plate 1 (see § 5.2.2). The incompleteness is once again apparent. Curiously, some Abell clusters do not seem to correspond to any enhancement in the density field†!

5.3.4 The Galaxy-Cluster Correlation Function

We present results for the galaxy-cluster angular correlation function, hereafter denoted $w_{g-cl}(\theta)$. This quantity is in essence a measure of the average radial profile† of the clusters. Measuring $w_{g-cl}(\theta)$ is important for two main reasons:

† See the top right corner of Plate 2, for example.

† Having nothing to do with the $\mathcal{L}$ profiles discussed in the previous section!

firstly, it tells us something about the dynamics of the clusters, and secondly, it is a quantity necessary in certain theoretical investigations, such as that of the completeness of the Abell catalogue by Dekel *et al.* (1989). The best previous measurement of this quantity is that of Lilje and Efstathiou (1988), which used the Lick counts. The problems associated with this catalogue have been discussed in Chapter 2, so we shall not dwell on them further here.

Because the number of galaxies in the catalogue is large, it is possible to get a robust estimate for $w_{g-cl}(\theta)$ even with a small number of clusters (whereas the cluster-cluster correlation function is very noisy, having only $\frac{1}{2}N_{cl}(N_{cl}-1)$ pairs).

Figure 12 shows the galaxy-cluster angular correlation function. The squares represent the values for the north field, and the triangles, the south. The open symbols are for the sample of clusters having $f \geq 1.5$, which is given in Tables 3 and 4. The filled symbols are for the sample with $f \geq 1.9$, which contains 39 clusters in the north and 52 in the south.

The amplitude of $w_{g-cl}(\theta)$ is very strong at least to 2°. If the average cluster is at a redshift of 0.1, this indicates that we are easily detecting the halo of the clusters out to $10h^{-1}$ Mpc from the cluster centers.

We fit two power laws to the data, one at small angles ($0.025° \leq \theta \leq 0.2°$), and one at large angles ($0.2° \leq \theta \leq 10°$). If we express the correlation function as $w_{g-cl}(\theta) \approx \theta^{-\gamma}$, we find, for the north, $\gamma = 0.75$ for small angles, and $\gamma = 0.83$ for large angles. For the south field, we obtain $\gamma = 0.96$ for small angles, and $\gamma = 1.00$ for large angles.

It is unclear how the difference between the filled and unfilled symbols should be interpreted. If the clusters with $f \geq 1.9$ are systematically more distant than those with $f \geq 1.5$, then we would expect the filled symbols to lie to the top and right of the open symbols in Figure 12, as in the scaling of $w(\theta)$ for galaxies in Chapter 3. If the distance–f correlation is weak, however (as this author suspects it is), the only effect would be to move the filled symbols upwards, *i.e.* , only a richness effect is seen. The data seem to favour this latter hypothesis, but a very good fit could be obtained either way. In any case it is reassuring that both samples show the

same behaviour; this may indicate that we are detecting fundamentally the same physical systems at both values of f.

Note the presence of a break at $\theta \approx 0.2°$. This kink is actually predicted by the standard biased CDM. On Figure 13, we reproduce (with the author's permission) a figure taken from the paper of Lilje and Efstathiou (1988), which shows the galaxy-cluster correlation function expected from a standard CDM model. The absence of redshifts makes it difficult to make quantitative comparisons with this model; however, the correspondence in the shapes of the models to our measurements is almost unbelievably good. In particular, if we make the gross assumption that all clusters are at a redshift of 0.1, *both* features, the bend at $0.5h^{-1}$ Mpc and the bend at $1.5h^{-1}$ Mpc are observed.

This is remarkable, but unfortunately is not a very strong test of CDM, which reproduces the power spectrum on small scales quite well, but has problems on larger scales.

5.4 Conclusion

From a highly uniform catalogue of galaxies, we have compiled a sample of clusters of galaxies using an objective algorithm. We believe this catalogue to be suitable for statistical analyses. Its main deficiency is its small size, and lack of a distance indicator. We are working on obtaining redshifts for these clusters. with which we will attempt to measure the spatial correlation function $\xi(r)$ for clusters.

An analysis of the multiplicity function makes it seem unlikely that we can obtain more reliable catalogues than this one using only 2-d information. A severe limitation at present is the limited dynamic range of the catalogue, rendering the magnitude information virtually useless.

Comparison of this objectively selected catalogue of clusters of galaxies with that of Abell's reveals flaws in the Abell catalogue which render the previous statistical work on it suspect. In particular, clusters belonging to the same distance and richness classes have a large dispersion of f values. It is also shown that Abell did not do his background correction sufficiently accurately.

The choice of whether to use a local or global background correction is unclear. We prefer using the local background because *i*) it yields multiplicity function which agree, so don't have to invoke amazing large-scale structure, and *ii)* we know that the universe really composed of clusters superposed on a uniform field, and so the local background may be more robust (it is easy to imagine how a background "sheet" of galaxy might cause an algorithm which uses global subtraction to detect a great number of small, spurious enhancements as clusters).

The lack of redshifts for our clusters prevents a detailed comparison of our measurement of $w_{g-cl}(\theta)$ with theory. Qualitatively, however, CDM is able to produce the observed shape of $w_{g-cl}(\theta)$ quite accurately. Assuming reasonable redshifts for our clusters ($z \geq 0.1$), the clusters are easily detected out to radial distances of $15h^{-1}$ Mpc.

Table 3 North Objective Clusters

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
1	14 30 57.1	31 47 52	1.603	5799	5258	4856	4565	4296	4048	3871	3729	3552	3385
2	14 31 44.5	29 39 16	1.632	5906	5191	4657	4250	3952	3646	3425	3275	3088	2938
3	14 35 50.3	30 20 17	1.715	—	—	—	3914	3556	3264	3060	2921	2762	2644
4	14 35 57.5	30 35 50	1.695	6132	5298	4733	4313	3989	—	3510	3366	3187	3061
5	14 36 55.8	30 56 19	1.588	5744	4958	4419	3984	3635	3314	3112	2939	2719	2583
6	14 39 15.8	28 47 25	1.702	—	—	—	—	—	3090	2876	2753	2561	2460
7	14 40 11.4	28 41 38	1.577	5706	5061	4610	4219	3922	3663	3451	3313	3117	2991
8	14 40 16.4	29 56 48	1.537	5561	4780	4231	3804	3455	3117	2873	2700	2491	2330
9	14 41 8.6	31 53 6	1.880	6803	5972	5347	4844	4475	4103	3830	3639	3394	3209
10	14 41 22.4	31 9 2	1.621	5864	5133	4622	4207	3858	3573	3355	3175	2976	2841
11	14 42 49.5	31 24 1	2.006	—	—	5215	4722	4366	4041	3819	3653	3437	3287
12	14 43 14.2	31 39 59	1.923	6956	6002	5401	4939	4567	4217	3949	3791	3562	3421
13	14 43 18.7	30 11 32	1.585	5734	4993	4471	4074	3733	3440	3229	3067	2856	2707
14	14 44 23.4	30 20 40	1.501	—	—	3903	3465	3147	2858	2644	2513	2328	2193
15	14 45 1.5	29 52 41	1.631	5902	5104	4560	4153	3806	3492	3265	3097	2864	2732
16	14 45 23.5	30 39 36	1.505	5444	4732	3917	3487	3489	2908	2998	2837	2647	2511
17	14 45 52.1	30 12 53	1.558	5638	4822	4244	3802	3426	3116	2887	2706	2489	2334
18	14 46 59.2	30 45 17	1.633	5910	5068	4495	4054	3697	3353	3117	2908	2701	2541
19	14 47 6.4	33 46 57	1.640	—	—	—	—	3361	3032	2792	2664	2435	2330
20	14 48 20.9	33 59 30	1.626	5882	5155	4655	4237	3919	3624	3399	3262	3066	2928
21	14 48 38.7	28 4 47	1.523	5510	4728	4166	3723	3378	3077	2839	2701	2499	2368
22	14 49 12.2	30 23 30	1.657	—	—	—	—	—	3009	2825	2698	2515	2391
23	14 49 13.4	31 30 40	2.107	7625	6763	6182	5712	5359	5002	4753	4602	4383	4225
24	14 49 20.0	30 52 21	2.163	7828	6809	6141	5555	5099	4703	4358	4151	3906	3681
25	14 49 34.0	30 35 33	1.772	—	—	—	4045	3689	3382	3142	2984	2760	2646
26	14 50 33.4	28 8 21	2.074	7505	6605	5921	5403	4968	4587	4304	4100	3813	3601
27	14 52 5.4	28 19 28	1.974	7143	6218	5596	5152	4761	4455	4181	4021	3826	3653
28	14 54 4.5	23 49 33	2.023	7321	6199	5537	4797	4479	4094	3793	3611	3334	3178
29	14 54 53.9	24 54 12	1.550	5609	4951	4484	4131	3848	3583	3394	3254	3080	2945
30	14 55 15.0	25 35 1	1.570	5681	4906	4219	3676	3292	2994	2914	2624	2564	2448

Table 3 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
31	14 56 0.2	23 16 0	1.799	—	—	—	—	—	—	2976	2821	2620	2503
32	14 56 10.1	23 35 19	1.914	6924	5937	5259	4715	4316	3930	3658	3457	3212	3036
33	14 56 31.8	25 7 47	1.788	6468	5727	5190	4744	4386	4030	3793	3630	3394	3252
34	14 56 45.1	28 1 6	2.127	7695	6835	6185	5692	5310	4919	4631	4451	4213	4034
35	14 57 37.1	24 17 52	1.828	6614	5781	5219	4749	4395	4033	3772	3614	3366	3232
36	14 58 13.9	23 9 8	1.886	6822	5920	5265	4735	4351	3989	3711	3514	3256	3068
37	14 59 47.6	23 1 54	1.647	—	—	4282	3826	3566	3274	3054	2901	2751	2611
38	14 59 49.6	24 52 44	1.620	5860	5161	4641	4226	3916	3599	3381	3211	2967	2846
39	15 0 30.5	27 6 56	1.540	5571	4773	5105	4616	4223	3886	3596	3454	3221	3068
40	15 1 20.2	23 15 47	2.338	8461	7370	6584	5990	5424	4904	4620	4405	4094	3918
41	15 1 27.7	35 59 55	1.737	6284	5388	4804	4336	3980	3616	3326	3181	2975	2779
42	15 1 30.0	27 21 35	2.128	7700	6801	6150	5606	5229	4836	4541	4322	4079	3835
43	15 2 20.4	28 40 13	2.033	7357	6270	5546	4996	4562	4223	3949	3755	3097	3369
44	15 2 21.6	23 3 33	2.245	—	—	5835	5315	4827	4412	4118	3898	3626	3436
45	15 2 49.3	35 36 57	1.584	—	—	4118	3747	3435	3141	2973	2836	2650	2550
46	15 2 52.7	26 11 6	1.741	6300	5563	5039	4623	4290	3942	3710	3543	3320	3136
47	15 3 7.3	36 8 42	1.812	6556	5671	5053	4572	4218	3836	3587	3398	3128	2992
48	15 3 10.9	23 23 11	2.017	—	—	5242	4671	4258	3834	3545	3332	3080	2860
49	15 3 21.0	35 50 39	1.625	—	—	—	—	3331	—	2846	2706	2502	2379
50	15 3 46.6	23 34 51	1.569	—	—	—	—	—	—	—	2440	2263	2141
51	15 3 59.4	27 30 48	2.245	8121	6889	6357	5525	5281	4605	4534	4322	3735	3810
52	15 4 26.1	27 11 50	1.644	—	—	—	—	—	—	2726	2594	2314	—
53	15 4 35.7	23 14 57	1.658	—	—	—	—	—	—	—	2578	2350	2252
54	15 4 53.2	27 57 2	1.520	—	—	—	—	—	—	—	2364	2209	2070
55	15 4 58.4	25 50 40	1.673	6051	5243	4636	4197	3804	3476	3223	3059	2825	2667
56	15 5 43.0	26 34 46	1.628	5891	5071	4529	4098	3807	3503	3295	3153	2976	2856
57	15 6 52.2	27 17 52	2.214	8011	7039	6370	5826	5371	4957	4605	4405	4071	3855
58	15 6 58.5	28 21 50	1.528	—	—	3972	3560	3275	2991	2801	2667	2494	2367
59	15 6 59.7	27 48 9	1.612	—	—	—	—	—	—	2667	2518	2294	2178
60	15 7 8.9	26 42 16	1.885	6822	5962	5355	4888	4490	4161	3873	3720	3492	3333

Table 3 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
61	15 7 12.0	30 23 55	1.680	6079	5424	4992	4609	4321	4024	3822	3671	3446	3304
62	15 7 43.2	28 3 43	1.678	6073	5250	4661	4211	3838	3513	3275	3105	2872	2721
63	15 8 16.4	33 38 32	2.680	9696	8742	8079	7512	7043	6545	6212	5981	5592	5395
64	15 8 28.8	36 52 9	2.172	—	—	—	4957	4530	4105	3782	3590	3307	3108
65	15 9 35.7	36 40 41	2.012	7279	6266	5611	5111	4798	4410	4169	4027	3817	3630
66	15 9 37.6	37 8 10	1.831	—	—	—	4179	3763	3413	3138	2990	2774	2615
67	15 9 47.6	28 17 1	1.578	—	—	—	—	—	2865	2680	2524	2350	2232
68	15 9 51.5	34 54 33	1.639	5931	5114	4554	4095	3707	3403	3165	3008	2817	2629
69	15 10 6.2	35 8 16	1.771	—	—	—	—	—	—	2930	2795	2624	2463
70	15 11 36.0	37 19 17	2.046	7402	6454	5772	5268	4818	4376	4092	3887	3569	3431
71	15 11 38.7	28 11 19	1.804	6527	5665	5067	4608	4200	3841	3566	3389	3143	2961
72	15 11 50.3	24 22 52	1.588	5746	4993	4433	3996	3648	3325	3085	2939	2725	2586
73	15 12 51.9	37 3 48	1.672	6048	5421	4690	4012	3667	3279	3048	2842	2597	2444
74	15 13 2.6	25 2 25	1.754	6345	5534	4945	4470	4151	3798	3564	3400	3163	3002
75	15 13 3.8	29 2 35	1.633	5907	5187	4678	4279	3953	3633	3421	3285	3056	2953
76	15 14 18.5	23 43 42	1.508	5455	4860	4449	4122	3841	3602	3414	3276	3090	2989
77	15 14 48.1	37 18 5	1.607	—	—	—	3668	3364	3430	3190	3071	2887	2755
78	15 14 49.6	36 17 52	1.582	5723	4943	4420	3980	3670	3346	3094	2963	2784	2615
79	15 15 12.1	33 28 26	1.635	5914	5192	4686	4256	3918	3584	3308	3152	2923	2743
80	15 15 21.3	28 52 26	1.576	5702	4850	4242	3772	3414	3077	2837	2684	2479	2331
81	15 15 30.3	29 6 21	1.582	—	—	—	—	—	—	2617	2476	2264	2138
82	15 15 33.6	28 1 28	1.572	5687	4982	4490	4109	3814	3517	3317	3164	2970	2827
83	15 16 16.8	36 38 11	1.516	5485	4773	4242	3845	3513	3208	2946	2832	2644	2472
84	15 16 40.4	28 48 17	1.651	—	—	4292	3834	3486	3163	2948	2784	2571	2421
85	15 17 11.9	28 29 6	2.110	7633	6618	5878	5290	4860	4416	4096	3904	3602	3406
86	15 17 27.8	27 45 6	1.637	—	4943	4369	3896	3544	3194	2883	2776	2513	2383
87	15 18 23.3	29 1 29	2.221	8036	7002	6307	5757	5339	4878	4585	4407	4143	3945
88	15 18 46.5	28 52 53	1.738	—	5248	—	—	3797	3393	3112	—	—	—
89	15 19 19.2	29 35 30	1.845	6675	5860	5286	4833	4487	4149	3911	3746	3505	3332
90	15 19 20.5	30 47 10	3.074	11120	10056	9309	8624	8174	7577	7199	6948	6593	6303

Table 3 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
91	15 19 31.3	28 43 1	1.799	6508	5566	4933	4428	4015	3678	3423	3254	2999	2863
92	15 19 34.0	28 16 58	2.763	9995	8776	7928	7214	6713	6164	5795	5523	5161	4892
93	15 19 35.1	32 13 51	2.102	7606	6745	6142	5667	5249	4872	4604	4397	4117	3946
94	15 19 52.6	29 16 6	1.788	—	5397	—	4287	—	3517	3219	3057	2859	2672
95	15 20 11.1	29 13 42	1.739	6291	5467	4737	—	—	—	—	3057	—	—
96	15 20 28.9	27 51 18	3.183	11518	10206	9327	8552	8001	7436	6971	6708	6370	6037
97	15 20 32.0	31 38 9	1.848	—	—	—	—	3787	3493	3266	3111	2881	2738
98	15 21 17.5	31 7 54	2.431	8796	7752	7007	6435	5952	5505	5180	4948	4594	4401
99	15 21 25.6	28 26 25	2.007	—	6060	5333	4745	4321	3952	3676	3457	3216	3060
100	15 21 41.1	28 47 13	1.902	6882	5913	5260	4726	4368	3991	3753	3555	3330	3122
101	15 22 8.5	30 1 32	3.044	11013	9795	8893	8204	7588	7055	6659	6340	5959	5658
102	15 22 11.9	27 55 18	1.989	—	—	5171	4615	4227	3831	3590	3410	3166	2996
103	15 22 29.0	35 23 25	1.551	—	4684	4124	3709	3403	3099	2900	2770	2580	2424
104	15 23 5.8	37 7 20	1.839	6653	6022	5530	5097	4780	4424	4148	3955	3741	3530
105	15 23 35.8	29 48 51	2.142	—	—	—	—	—	3915	3620	3331	3132	2935
106	15 23 36.1	28 36 30	2.215	8015	6954	6217	5685	5211	4810	4527	4320	4029	3820
107	15 24 17.0	35 34 57	1.606	5811	4934	4377	3904	3560	3218	2973	2817	2606	2445
108	15 24 18.9	30 1 35	2.022	—	—	—	—	—	—	3345	3178	2855	2722
109	15 24 27.3	29 46 17	1.767	6392	5437	4832	4294	4119	3911	3377	3331	2964	2849
110	15 25 8.5	34 56 11	1.998	7227	6329	5668	5115	4680	4251	3937	3698	3385	3065
111	15 25 45.8	28 59 57	2.403	8693	7833	7144	6598	6178	5731	5373	5158	4832	4579
112	15 26 45.9	29 13 13	2.060	—	—	—	—	—	—	3408	3232	2995	2840
113	15 27 38.5	30 50 25	2.026	7331	6480	5840	5341	4939	4522	4239	4038	3755	3582
114	15 27 49.8	29 9 42	1.962	—	5925	5277	4761	4407	4041	3805	3648	3429	3277
115	15 28 46.9	28 5 36	1.632	5904	5075	4498	4079	3736	3429	3227	3070	2870	2725
116	15 30 33.5	28 11 19	1.853	6704	5912	5377	4928	4566	4236	3963	3804	3540	3389
117	15 31 19.0	31 16 45	1.534	5551	4871	4418	4047	3727	3473	3236	3120	2948	2825
118	15 32 19.3	29 13 47	1.554	—	—	—	—	—	2821	2649	2528	2368	2267
119	15 34 4.7	25 42 39	1.554	5622	4879	4330	3912	3545	3191	2918	2767	2577	2373

Table 4 South Objective Clusters

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
1	23 53 15.1	1 29 33	1.846	4600	4139	3803	3498	3282	3048	2890	2766	2597	2474
2	23 53 32.3	10 59 31	2.100	5233	4609	4137	3746	3408	3063	2871	2746	2518	2352
3	23 53 42.8	5 9 27	1.593	3970	—	—	—	—	2223	—	—	1807	1711
4	23 53 47.7	5 45 3	1.562	3892	3407	3046	2765	2543	2335	2198	2099	1961	1886
5	23 54 4.8	4 50 2	1.611	—	—	2923	2647	2401	2198	2037	1933	1798	1695
6	23 55 11.8	5 12 6	1.644	4097	3592	3237	2967	2757	2553	2414	2309	2162	2059
7	23 55 37.9	1 1 46	1.771	4415	3875	3492	3167	2940	2685	2495	2381	2202	2099
8	23 56 7.2	3 47 50	1.713	4269	3769	3416	—	2911	2693	2520	2412	2255	2154
9	23 57 9.6	4 41 0	1.547	3856	3313	2950	2678	2461	2272	2130	2028	1888	1800
10	23 57 26.7	5 22 38	1.656	—	—	—	—	—	2109	—	—	1711	1620
11	23 57 45.3	0 23 57	1.923	4793	4268	3927	3625	3410	3191	3003	2899	2747	2626
12	23 57 48.0	4 59 52	1.952	4865	4189	3712	3379	3098	2852	2671	2551	2400	2267
13	23 58 9.7	6 12 17	1.741	4339	3808	3388	3054	2826	2580	2409	2313	2156	2022
14	23 58 25.1	6 36 26	1.723	—	3615	—	2934	2674	2457	2277	2180	2029	1923
15	23 58 27.8	5 15 14	2.078	—	—	—	—	—	2647	2456	2332	2149	2022
16	23 58 47.0	5 55 42	1.727	4303	3760	3393	3079	2851	2637	2484	2359	2216	2119
17	23 59 7.0	5 29 16	1.959	—	—	—	—	—	2495	2313	—	1998	1891
18	23 59 56.7	12 24 59	1.751	—	3674	3324	3086	2883	2645	2508	2378	2228	2130
19	0 0 1.0	5 42 29	1.896	4726	4089	3638	3294	3006	2772	2561	2437	2281	2170
20	0 1 1.1	0 47 53	1.664	4146	3628	3248	2952	2720	2511	2349	2236	2107	1775
21	0 1 4.1	4 19 59	2.022	5038	4432	4019	3687	3443	3186	3001	2887	2736	2622
22	0 1 7.4	5 0 36	1.937	4826	4231	3825	3484	3228	2972	2793	2663	2480	2350
23	0 1 25.5	3 22 54	1.696	4227	3622	3189	2828	2572	2307	2114	—	—	1729
24	0 1 34.2	1 46 10	2.042	5090	4520	4066	3701	3423	3161	2950	2809	2596	2463
25	0 2 8.8	3 30 45	1.692	—	—	—	—	—	—	—	1854	1716	1623
26	0 2 12.2	5 9 20	1.576	—	—	—	—	—	—	—	1726	1593	1490
27	0 2 52.2	11 18 3	1.515	3775	3283	2941	2653	2436	2240	2086	1989	1861	1738
28	0 3 39.4	10 46 58	1.772	4415	3910	3536	3283	3041	2849	2691	2591	2421	2344
29	0 4 5.7	10 29 11	1.913	4767	4179	3766	3442	3176	2943	2769	2642	2455	2344
30	0 4 56.8	10 15 56	1.546	—	—	—	—	—	1969	1839	1753	1646	1579

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
31	0 5 17.0	−1 3 35	1.543	3845	3388	3064	2809	2607	2428	2287	2192	2075	1996
32	0 7 1.9	6 28 1	1.724	4295	3763	3371	3095	2862	2640	2494	2407	2274	2175
33	0 7 4.6	3 41 26	2.158	5377	4873	4475	4135	3912	3625	3450	3318	3126	3002
34	0 7 13.4	12 10 21	1.603	—	—	—	—	—	2042	1906	1804	1689	1591
35	0 7 51.2	5 47 53	1.580	—	—	—	—	—	2012	1867	1768	1632	1542
36	0 7 54.6	11 55 34	1.560	3887	3354	3002	2741	2544	2367	2232	2143	2040	1944
37	0 7 58.3	6 21 3	1.562	—	—	—	—	2234	2039	1915	1830	1700	1633
38	0 8 28.9	11 11 17	1.755	4373	3833	3444	3135	2904	2679	2512	2402	2231	2129
39	0 8 29.3	12 23 5	1.685	—	3535	3162	2871	2617	2421	2266	2172	2027	1931
40	0 8 46.0	12 8 6	1.640	—	3441	3018	2694	2494	2283	2150	2072	1958	1865
41	0 8 46.8	5 53 25	1.726	4302	3787	3441	3130	2896	2701	2549	2436	2292	2198
42	0 9 20.0	10 57 49	1.745	—	—	—	—	2497	2280	2140	2027	1882	1812
43	0 10 27.0	0 1 22	1.527	3805	3306	2954	2688	2496	2289	2156	2068	1938	1847
44	0 10 28.4	0 24 2	1.692	4216	3664	3272	2963	2723	2498	2339	2226	2060	1959
45	0 10 29.4	−1 50 15	1.640	4088	3599	3231	2966	2749	2545	2380	2294	2159	2079
46	0 11 0.5	0 38 1	1.624	—	—	—	—	—	—	1891	1804	1670	1572
47	0 11 2.1	10 24 3	1.558	—	—	—	2485	2268	2074	1922	1836	1705	1602
48	0 11 6.8	0 −14 37	1.617	—	—	—	—	—	—	1883	1786	1658	1567
49	0 11 17.4	10 46 14	1.853	4618	4094	3741	3428	3172	2936	2783	2645	2447	2334
50	0 11 20.6	0 −34 57	1.595	—	—	—	—	2282	2077	1944	1845	1719	1628
51	0 11 47.8	4 47 18	1.559	3886	3399	3059	2789	2584	2376	2233	2130	2000	1915
52	0 12 31.2	0 −28 48	1.931	4812	4150	3698	3313	3032	2782	2585	2455	2290	2162
53	0 13 45.7	2 26 9	1.680	4187	3703	3344	3061	2820	2616	2443	2338	2192	2087
54	0 14 0.5	0 −21 40	1.542	—	—	—	—	2206	2012	1875	1776	1648	1564
55	0 14 1.8	6 1 17	1.749	—	—	—	2790	2551	2364	2199	2121	1989	1877
56	0 14 17.6	0 −45 0	1.870	4661	4076	3681	3406	3154	2949	2790	2685	2568	2455
57	0 14 21.9	−1 22 10	2.193	5465	4878	4445	4091	3852	3586	3397	3249	3059	2942
58	0 14 24.7	8 22 47	1.680	4188	3678	3322	3029	2796	2569	2392	2288	2147	2037
59	0 14 34.7	6 29 57	2.399	5979	5302	4790	4406	4083	3747	3570	3380	3200	3020
60	0 15 5.6	−1 40 1	1.928	—	4044	—	3268	3008	2753	2558	2442	2303	2159

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
61	0 15 8.1	1 38 18	1.726	4301	3685	3254	2904	2535	2347	2121	2017	1856	1757
62	0 15 17.3	−1 6 41	1.948	4855	4213	3774	3396	3114	2826	2613	2469	2085	1981
63	0 15 31.3	0 6 37	1.769	4409	3837	3443	3117	2852	2623	2430	2314	2135	2032
64	0 15 42.4	6 29 36	1.733	—	—	—	—	2479	2275	2095	2008	1882	1784
65	0 15 54.0	1 28 46	1.751	—	—	—	—	—	2231	—	1959	1805	1718
66	0 16 14.9	0 −27 3	1.725	—	—	—	2751	2517	2316	2151	2067	1927	1843
67	0 16 17.8	1 12 26	1.914	4771	4113	3673	3315	3030	2789	2592	2467	2272	2157
68	0 16 32.8	−1 33 33	1.647	—	—	2989	2668	2437	2214	2060	1951	1789	1702
69	0 17 11.2	0 8 4	1.641	—	—	—	2617	2384	2149	2010	1888	1760	1655
70	0 17 13.7	−1 8 55	1.685	—	—	—	2688	—	2225	2069	2001	1805	1698
71	0 17 14.0	1 29 54	1.518	—	—	2755	2485	2270	2089	1974	1869	1765	1661
72	0 17 29.7	1 3 4	1.750	4361	3783	3345	3017	2753	2423	2279	2173	2043	1995
73	0 17 34.8	0 35 4	1.596	3978	3408	3006	2695	2466	2231	2065	1960	1821	1729
74	0 17 51.3	0 22 43	1.616	—	—	—	—	—	—	1881	1802	1667	1588
75	0 17 51.9	0 −46 6	2.035	—	—	—	—	—	—	2369	2237	—	1939
76	0 17 55.3	0 −33 36	2.018	5030	4318	3821	3429	3130	2865	2651	2536	2361	2251
77	0 17 57.6	0 −10 6	2.259	5630	4991	4524	4177	3883	3605	3382	3246	3035	2886
78	0 18 49.0	6 39 41	1.811	4514	4039	3682	3413	3163	2968	2787	2674	2525	2410
79	0 19 4.8	−1 8 39	3.283	8183	7197	6481	5888	5438	4962	4672	4443	4099	3881
80	0 19 20.6	0 −30 39	1.789	—	—	—	2853	—	2389	2226	2117	1980	1859
81	0 19 54.0	6 24 9	1.541	—	—	—	2459	2210	2018	1882	1775	1651	1559
82	0 19 55.1	−1 33 51	1.673	—	3510	3078	2735	2519	2312	2165	2067	1933	1856
83	0 20 27.8	6 51 7	1.518	3783	3289	2951	2686	2461	2278	2135	2029	1897	1819
84	0 20 35.9	0 −24 23	2.426	6046	5372	4904	4538	4243	3985	3779	3634	3452	3306
85	0 20 38.3	0 7 36	2.049	5106	4394	3884	3491	3204	2915	2690	2577	2368	2254
86	0 20 46.1	0 54 36	1.551	3865	3349	2990	2710	2483	2267	2108	1988	1838	1738
87	0 20 56.1	0 −7 14	2.106	—	—	—	—	3013	—	2477	2363	2163	2061
88	0 21 0.0	−2 0 44	2.467	6148	5404	4866	4421	4070	3709	3429	3257	2999	2833
89	0 21 20.9	0 −45 55	1.789	4459	—	3388	3071	2804	2595	2420	2320	2180	2068
90	0 21 23.2	−1 17 15	1.772	—	—	—	—	2563	2313	2149	2029	1866	1634

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
91	0 22 46.8	1 2 40	1.687	4211	3539	3121	2804	2580	2363	2206	1998	1962	1850
92	0 23 31.0	2 1 27	1.539	—	—	2792	2508	2309	2109	1966	1871	1746	1654
93	0 23 41.3	1 36 47	1.558	—	—	—	—	—	—	—	1707	1573	1494
94	0 24 1.0	1 20 26	1.985	4948	4338	3929	3572	3284	3042	2830	2707	2528	2398
95	0 24 3.3	0 54 45	1.775	4423	3844	3451	3139	2909	2694	2535	2428	2275	2173
96	0 24 20.7	0 15 14	1.872	4665	4038	3597	3248	2992	2710	2523	2391	2216	2094
97	0 24 23.6	1 12 44	1.700	—	—	—	—	2482	2175	1992	1863	1770	1689
98	0 24 46.4	2 7 43	1.889	4708	4132	3746	3449	3223	2987	2839	2729	2584	2473
99	0 25 28.2	1 55 30	1.534	—	—	2911	—	—	—	—	1681	1540	1448
100	0 26 16.1	0 24 7	1.693	—	3552	3158	2875	2649	2431	2265	2167	2034	1943
101	0 26 18.5	7 27 55	1.737	4330	3809	3450	3162	2950	2712	2560	2459	2331	2207
102	0 26 28.7	11 59 2	1.804	4496	3897	3489	3147	2886	2626	2440	2320	2120	2007
103	0 26 31.0	0 1 6	1.530	—	—	—	2440	2246	2035	1890	1810	1678	1600
104	0 26 41.3	11 3 42	1.522	3794	3322	2977	2698	2466	2263	2105	2018	1871	1778
105	0 26 47.0	0 40 25	1.547	—	—	—	—	—	—	—	1694	1556	1468
106	0 26 53.5	5 41 2	1.539	3836	3334	2978	2722	2497	2306	2166	2058	1910	1823
107	0 26 54.6	0 −27 44	2.877	7169	6309	5671	5170	4773	4374	4107	3886	3631	3461
108	0 27 13.6	11 39 48	2.116	5272	4625	4134	3770	3472	3187	2946	2825	2637	2491
109	0 27 45.3	6 41 27	1.539	3836	3349	3018	2739	2529	2315	2172	2079	1935	1854
110	0 28 46.8	6 7 12	1.589	3959	3451	3070	2770	2558	2335	2186	2090	1933	1845
111	0 28 47.5	4 54 23	1.576	3927	3360	2948	2659	2427	2226	2086	1990	1860	1760
112	0 28 58.0	5 10 12	1.591	—	—	—	—	—	—	1853	1787	1661	1559
113	0 29 3.2	2 9 33	2.175	5420	4772	4321	3937	3664	3388	3164	3021	2842	2704
114	0 29 8.9	12 8 45	1.660	4137	3685	3351	3089	2878	2674	2509	2409	2260	2167
115	0 29 29.5	0 0 −26	1.768	4405	3867	3489	3175	2948	2718	2542	2438	2285	2158
116	0 29 29.9	0 −57 55	1.809	4508	3950	3544	3216	2981	2746	2584	2457	2284	2183
117	0 29 38.1	2 18 20	2.353	—	—	—	—	—	—	—	2578	2400	2273
118	0 29 55.9	1 2 0	2.221	5535	4854	4383	3983	3730	3437	3234	3103	2925	2809
119	0 30 8.9	4 44 25	1.793	4469	3849	3413	3073	2805	2582	2417	2283	2133	2004
120	0 30 17.0	1 28 48	1.540	3838	3256	2864	2577	2338	2131	1972	1883	1743	1665

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
121	0 30 31.4	4 59 2	1.931	—	—	—	—	—	—	2248	2119	1954	1848
122	0 30 34.9	−2 6 29	1.768	4405	3879	3542	3256	3047	2831	2680	2579	2433	2336
123	0 30 40.6	3 39 13	1.860	4635	4050	3663	3321	3102	2858	2691	2572	2400	2306
124	0 31 11.2	1 2 23	1.934	—	—	3509	3172	2949	2716	2553	2458	2326	2223
125	0 31 22.7	0 −37 15	1.885	4697	4112	3685	3355	3094	2838	2651	2521	2113	2007
126	0 31 22.7	5 14 27	1.743	4344	3708	3302	2981	2744	2548	2396	2296	2143	2055
127	0 31 30.1	6 0 5	1.885	4699	4102	3675	3329	3077	2825	2599	2487	2313	2228
128	0 31 32.3	−1 3 55	1.707	4253	3741	3365	3096	2864	2642	2483	2365	2213	2107
129	0 31 46.5	2 6 58	1.882	4690	4068	3663	3336	3057	2849	2688	2574	2400	2299
130	0 31 46.6	0 33 21	2.229	5555	4927	4454	4099	3793	3520	3305	3173	2978	2833
131	0 31 54.1	0 7 31	1.714	4271	3673	3252	2954	2701	2481	2317	2209	2067	1951
132	0 31 55.9	−2 13 45	2.531	—	—	—	—	3621	3379	3200	3064	2885	2759
133	0 32 18.8	11 57 31	1.711	4265	3740	3355	3064	2819	2602	2451	2337	2182	2070
134	0 32 35.6	5 13 39	1.506	—	—	—	—	—	1918	1770	1687	1548	1471
135	0 32 46.8	0 46 25	1.539	—	—	—	—	—	—	1802	1686	1542	1442
136	0 32 56.2	1 50 52	1.748	—	—	—	—	—	2226	2070	1945	1812	1689
137	0 32 58.8	1 22 23	2.560	6380	5600	5034	4604	4215	3902	3656	3487	3233	3080
138	0 33 2.0	3 5 39	1.515	3776	3283	2941	2646	2437	2225	2053	1962	1815	1727
139	0 33 2.5	5 57 54	1.848	4606	3963	3550	3237	3016	2762	2601	2473	2322	2218
140	0 33 6.7	2 14 30	1.856	4624	3969	3535	3172	2902	2631	2446	2293	2147	2006
141	0 33 20.1	0 −22 45	1.735	4325	3733	3354	3050	2823	2613	2442	2343	2211	2094
142	0 33 21.7	−1 29 3	1.526	3802	3322	2984	2708	2480	2284	1936	2022	1891	1787
143	0 33 23.8	0 −1 26	1.851	4613	3934	3447	3081	2840	2575	2402	2269	2084	1975
144	0 33 24.5	1 5 11	1.681	—	—	—	—	—	—	—	1842	1686	1582
145	0 33 27.5	2 46 17	1.526	3804	3297	2954	2676	2467	2283	2141	2032	1898	1808
146	0 33 52.2	6 17 57	1.523	3795	3335	2997	2748	2534	2347	2206	2103	1983	1884
147	0 34 1.8	0 32 58	1.512	—	—	2744	2442	2205	1997	1832	1749	1606	1504
148	0 34 11.3	0 10 58	1.827	—	3833	—	3009	2749	2500	2310	2201	2048	1923
149	0 34 27.3	1 15 4	1.557	—	—	—	—	—	—	1812	1718	—	1503
150	0 34 38.6	2 35 49	1.562	—	—	—	—	2235	2051	1885	1813	1686	1602

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
151	0 34 39.5	−1 32 40	1.600	3987	3454	3094	2813	2574	2374	2215	2114	1961	1860
152	0 35 3.5	0 −11 26	2.192	5463	4767	4305	3927	3621	3330	3120	2977	2770	2624
153	0 35 23.3	1 17 18	1.514	3772	3235	2865	2561	2344	2131	1982	1879	1592	1503
154	0 35 35.6	−2 5 14	1.644	4096	3525	3139	2809	2572	2350	2174	2068	1925	1809
155	0 35 38.5	2 28 13	1.604	3998	3477	3108	2839	2613	2403	2268	2170	2054	1968
156	0 35 58.9	−1 17 18	1.690	4213	3628	3219	2877	2633	2374	2188	2073	1910	1810
157	0 36 7.2	3 38 6	1.544	3848	3289	2918	2621	2299	2165	1952	1895	1749	1667
158	0 36 18.7	0 −13 23	1.629	4280	3444	3026	2762	2470	2206	1942	1904	1611	1518
159	0 36 31.0	0 −32 14	1.726	—	3621	3204	2914	2653	2439	2278	2182	2026	1923
160	0 36 41.7	0 38 1	1.731	4314	3780	3409	3134	2890	2700	2547	2428	2298	2198
161	0 36 57.2	11 8 38	1.588	3959	3502	3162	2904	2708	2512	2375	2263	2157	2050
162	0 37 1.0	6 57 19	1.513	3770	3281	2927	2650	2383	—	2062	1974	1825	1748
163	0 37 2.7	−2 9 22	2.001	4987	4424	4014	3681	3434	3178	2997	2866	2704	2596
164	0 37 23.9	6 27 47	1.978	4930	4321	3893	3536	3267	2988	2777	2642	2491	2309
165	0 37 25.6	−1 9 27	1.528	—	—	2773	2497	2256	2053	1906	1814	1675	1620
166	0 37 38.7	0 −5 32	1.568	—	—	—	—	—	—	—	1718	1574	1489
167	0 37 41.7	0 −23 1	1.917	4776	4161	3716	3367	3111	2864	2665	2553	2391	2258
168	0 37 58.2	0 32 42	1.574	3922	3414	3062	2767	2569	2341	2187	2092	1932	1852
169	0 38 24.9	2 15 23	1.791	4464	3867	3446	3120	2879	2646	2449	2353	2191	2072
170	0 38 54.6	2 29 38	1.698	—	—	—	—	—	2163	2012	1910	1765	1661
171	0 40 11.0	12 12 57	1.524	3798	3296	2930	2631	2418	2191	1992	1903	1758	1663
172	0 40 13.9	0 −36 18	1.512	3767	3277	2952	2688	2498	2298	2115	2039	1875	1793
173	0 40 20.6	1 21 12	1.569	—	3291	2882	2559	2345	2124	1971	1877	1734	1645
174	0 40 29.4	1 1 2	1.772	—	—	—	—	—	2257	2086	1969	1834	1747
175	0 40 31.2	0 −4 42	2.156	5373	4684	4173	3756	3452	3111	2864	2704	2503	2295
176	0 40 41.4	1 48 37	2.202	5487	4802	4302	3898	3605	3312	3090	2952	2766	2624
177	0 40 44.9	2 15 54	1.549	—	—	2811	2530	2321	2131	1974	1886	1768	1705
178	0 41 6.8	3 10 17	1.530	3812	3328	2962	2661	2438	2184	1987	1821	1714	1635
179	0 41 22.9	12 10 2	1.508	3758	3228	2851	2555	2314	2109	1964	1864	1718	1628
180	0 41 32.5	0 54 14	2.258	5627	4902	4385	3966	3612	3302	3042	2874	2694	2522

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
181	0 41 41.7	−1 55 51	1.538	3833	3307	2945	2671	2440	2239	2099	1997	1869	1769
182	0 41 48.2	0 14 16	1.909	—	—	—	—	2730	—	2266	2136	1975	1868
183	0 42 13.0	−1 12 59	1.812	—	—	—	—	—	2308	2140	2027	1878	1769
184	0 42 13.2	3 0 38	1.590	3963	3451	3104	2824	2627	2438	2291	2177	2046	1965
185	0 42 25.1	11 3 33	1.561	—	—	—	—	—	—	—	1710	1598	1501
186	0 42 30.8	−2 7 15	1.511	—	—	—	—	2162	1987	1855	1754	1623	1552
187	0 42 41.8	0 29 57	1.985	—	4164	—	—	2960	2712	2477	2362	2157	2036
188	0 42 44.8	1 28 41	2.110	5259	4545	4036	3624	3321	3005	2796	2635	2419	2292
189	0 42 47.9	9 45 45	1.761	4388	3917	3566	3272	3085	2859	2702	2602	2468	2335
190	0 43 2.5	−1 39 2	1.673	—	—	—	—	—	—	—	1833	1671	1580
191	0 43 42.1	−1 59 4	1.928	4806	4152	3695	3328	3051	2766	2545	2422	2259	2081
192	0 43 44.0	0 −15 55	2.437	6075	5433	4966	4590	4256	3966	3722	3571	3351	3178
193	0 44 6.4	1 16 21	1.567	—	3609	3167	2499	2621	2409	2227	2139	2007	1898
194	0 44 20.8	0 42 52	1.951	4862	4197	3724	3346	3028	2757	2554	2418	2104	2107
195	0 44 32.8	−1 9 16	3.095	7714	6556	5918	5286	4790	4325	4145	3936	3647	3485
196	0 44 40.5	0 6 43	1.830	4561	3976	3566	3222	2919	2695	2404	2297	2152	1983
197	0 45 3.3	0 38 57	1.739	—	—	—	—	2489	2219	2024	1908	1758	1637
198	0 45 15.5	0 −32 46	1.549	—	—	—	—	2216	2002	1855	1762	1617	1529
199	0 45 24.9	0 25 55	1.597	3981	3388	3010	2711	2494	2269	2118	2008	1869	1769
200	0 45 49.0	0 55 18	1.513	—	3331	2855	2489	2248	1959	1776	1666	1497	1395
201	0 45 52.7	1 5 16	2.172	5413	4831	4385	4053	3765	3533	3341	3213	3020	2853
202	0 45 58.5	−2 1 50	1.627	—	—	—	—	—	—	1894	1800	1657	1564
203	0 46 38.5	−1 53 16	1.643	4095	3548	3194	2905	2664	2476	2304	2219	2082	1984
204	0 46 53.6	1 34 33	1.659	4134	3635	3278	2990	2770	2585	2443	2342	2199	2108
205	0 46 54.1	1 50 1	1.667	—	—	—	—	—	—	1941	1858	1727	1647
206	0 49 34.1	−1 20 5	1.653	4119	3640	3313	3070	2858	2670	2514	2418	2267	2182
207	0 50 7.7	1 0 26	1.542	—	—	—	2460	2236	2036	1876	1791	1652	1556
208	0 50 9.0	0 40 19	1.665	4149	3629	3260	2934	2714	2463	2285	2163	1995	1870
209	0 50 15.1	5 41 57	2.055	5120	4552	4167	3843	3555	3301	3135	2980	2825	2690
210	0 50 40.3	11 12 47	1.824	4545	3908	3455	3096	2790	2530	2354	2234	2052	1941

Table 4 continued

N	$\alpha_{1950.00}$	$\delta_{1950.00}$	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
211	0 50 51.0	10 42 59	2.775	6915	5868	5193	4612	4216	3847	3360	3176	3117	2937
212	0 50 51.3	9 55 57	1.780	—	—	3229	2949	2679	2504	2338	2233	2118	2020

Table 5. North Abell Clusters

N	St	D	R	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
1929	Y	6	2	1.26	4573.4	3992.3	3615.2	3272.1	3056.8	2787.6	2615.2	2510.4	2354.8	2245.8
1930	Y	5	1	1.60	5783.3	5205.0	4831.4	4475.0	4253.0	3958.7	3773.0	3657.2	3488.8	3358.4
1934	Y	6	3	1.63	5915.2	5101.6	4565.9	4085.5	3787.6	3424.2	3195.3	3059.2	2856.3	2720.5
1941	Y	6	1	1.68	6070.5	5264.1	4746.8	4290.4	4015.7	3678.5	3472.5	3350.6	3171.6	3046.9
1943	Y	6	2	1.64	5923.1	5094.1	4534.9	4024.4	3697.3	3292.8	3031.2	2870.4	2626.5	2469.6
1944	Y	5	1	1.68	6070.5	5264.1	4746.8	4290.4	4015.7	3678.5	3472.5	3350.6	3171.6	3046.9
1952	Y	6	2	1.30	4704.7	4075.6	3678.1	3327.6	3116.2	2860.3	2702.2	2610.6	2476.5	2382.4
1954	Y	6	2	1.57	5692.8	5043.7	4616.0	4216.3	3964.6	3644.6	3441.7	3318.3	3137.2	3007.2
1956	Y	6	2	1.88	6785.9	5860.2	5241.1	4680.7	4328.8	3895.1	3620.1	3454.4	3206.0	3042.1
1957	Y	6	3	1.37	5051.0	4173.4	3603.7	3123.2	2833.7	2492.7	2283.7	2160.3	1973.3	1859.1
1958	Y	6	2	1.62	5874.6	5110.1	4609.4	4158.4	3878.3	3527.9	3308.0	3174.5	2976.0	2841.3
1961	Y	6	3	1.79	6482.5	5580.9	5008.3	4505.6	4200.6	3824.9	3589.4	3447.4	3237.1	3090.6
1963	Y	6	2	1.88	6811.3	5894.9	5292.9	4752.8	4417.9	4003.9	3744.9	3589.4	3358.3	3204.2
1965	N	6	0	1.27	4609.4	3957.8	3526.9	3141.6	2901.0	2605.4	2418.8	2306.3	2137.4	2026.7
1968	Y	6	1	1.16	4180.9	3603.0	3231.6	2903.1	2701.4	2454.0	2297.8	2204.2	2064.6	1968.7
1979	Y	5	2	2.00	7247.0	6322.7	5721.3	5177.7	4841.7	4425.1	4162.6	4005.5	3772.9	3610.5
1982	N	5	0	2.14	7760.4	6755.0	6094.5	5495.6	5122.3	4654.1	4358.2	4177.1	3906.4	3720.7
1984	Y	5	2	2.03	7356.5	6415.7	5803.9	5247.4	4898.0	4458.7	4176.6	4003.5	3745.3	3564.6
1987	N	6	0	1.36	4909.7	4270.7	3857.5	3486.3	3257.7	2970.0	2788.3	2677.3	2512.8	2397.7
1990	Y	5	3	1.99	7191.4	6200.0	5560.4	4994.9	4647.8	4219.4	3949.7	3787.0	3544.4	3377.0
2004	N	5	0	1.79	6469.8	5703.6	5197.2	4721.7	4416.4	4026.6	3772.8	3615.9	3382.9	3217.2
2005	Y	4	2	2.11	7650.3	6739.1	6142.7	5591.3	5246.5	4808.3	4528.8	4358.2	4106.8	3925.3
2008	Y	6	2	1.83	6618.4	5631.0	4977.3	4396.2	4031.4	3588.7	3306.6	3136.4	2879.9	2715.7
2017	Y	5	1	2.08	7529.7	6434.4	5712.1	5067.5	4664.1	4176.6	3866.0	3681.5	3405.5	3226.1
2019	N	4	0	2.01	7264.7	6222.9	5523.2	4891.8	4491.0	3999.2	3683.3	3490.9	3199.4	3012.1

Table 5. North Abell Clusters (continued)

N	St	D	R	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
2021	Y	5	1	2.26	8183.3	7009.1	6243.3	5565.5	5147.5	4645.0	4328.8	4142.7	3865.1	3681.0
2022	Y	3	1	1.91	6923.4	5922.8	5255.6	4655.5	4274.5	3805.9	3503.0	3317.7	3035.9	2853.4
2025	Y	5	1	1.34	4834.8	4235.0	3847.4	3498.8	3287.3	3023.9	2861.5	2765.3	2624.6	2524.8
2034	Y	5	2	2.59	9367.0	8404.7	7768.5	7157.6	6771.0	6265.7	5942.7	5743.8	5453.2	5234.3
2042	Y	5	1	1.92	6952.6	6072.8	5512.9	5012.0	4706.5	4327.2	4090.4	3949.5	3742.9	3593.9
2046	Y	5	1	1.46	5301.3	4429.5	3845.6	3336.0	3019.2	2645.9	2410.6	2272.3	2063.5	1937.8
2049	N	5	0	1.33	4813.0	4262.4	3908.1	3583.8	3385.8	3135.5	2980.3	2887.5	2752.7	2652.6
2056	Y	5	1	2.11	7632.8	6569.4	5862.8	5222.3	4816.6	4317.3	3995.5	3800.3	3506.4	3312.7
2059	Y	5	1	2.12	7676.7	6652.1	5994.9	5408.0	5045.3	4595.4	4308.0	4133.6	3874.0	3691.6
2061	Y	4	1	3.08	11139.6	10046.2	9326.5	8623.8	8174.4	7577.2	7190.2	6947.7	6592.6	6320.1
2062	Y	5	1	2.06	7440.1	6544.3	5952.5	5406.2	5063.7	4629.3	4353.9	4185.2	3935.1	3758.7
2065	Y	3	2	3.07	11111.6	9747.6	8841.9	8001.6	7467.6	6792.7	6356.3	6088.2	5688.3	5409.2
2067	Y	4	1	2.25	8128.6	7089.0	6414.6	5801.4	5416.5	4932.3	4622.2	4431.5	4146.9	3947.0
2069	Y	5	2	2.98	10765.7	9469.0	8613.4	7821.4	7322.3	6689.5	6283.3	6035.1	5665.8	5403.9
2070	Y	6	1	1.31	4722.4	4075.8	3664.3	3303.2	3083.8	2814.9	2648.7	2550.0	2404.0	2303.9
2071	Y	5	1	1.79	6484.3	5824.6	5373.5	4930.0	4639.4	4256.1	4004.3	3845.6	3609.5	3437.5
2073	Y	5	1	2.21	8010.8	6928.2	6213.1	5566.4	5160.1	4658.5	4339.3	4146.7	3858.9	3665.9
2079	Y	3	1	2.34	8483.3	7481.1	6816.2	6190.4	5786.8	5269.4	4929.9	4718.6	4402.0	4178.1
2083	Y	5	1	2.01	7288.6	6402.6	5815.7	5270.9	4924.9	4486.8	4204.8	4031.1	3772.3	3591.2
2089	Y	4	1	1.82	6595.1	5827.3	5323.0	4851.1	4552.1	4170.3	3923.7	3771.8	3546.1	3383.9
2092	Y	4	1	1.51	5449.8	4785.1	4358.7	3970.7	3730.7	3427.6	3236.1	3119.8	2948.2	2824.5

Table 6. South Abell Clusters

N	St	D	R	f	$\mathcal{L}(1)$	$\mathcal{L}(2)$	$\mathcal{L}(3)$	$\mathcal{L}(4)$	$\mathcal{L}(5)$	$\mathcal{L}(6)$	$\mathcal{L}(7)$	$\mathcal{L}(8)$	$\mathcal{L}(9)$	$\mathcal{L}(10)$
3	Y	5	1	1.95	4861.2	4329.0	3956.0	3638.8	3391.1	3144.8	2937.4	2807.5	2645.5	2505.0
4	Y	6	1	1.63	4073.1	3576.5	3228.4	2957.0	2755.5	2555.5	2409.9	2317.7	2189.4	2097.4
9	Y	6	1	1.43	3567.4	3181.2	2917.7	2716.3	2543.9	2397.7	2282.5	2199.6	2098.3	2012.0
16	Y	5	2	2.38	5926.5	5223.3	4716.8	4299.4	3982.8	3649.5	3412.1	3254.6	3026.1	2867.3
17	Y	6	1	1.61	4011.6	3474.6	3111.1	2775.2	2558.0	2287.3	2110.6	2001.6	1836.8	1728.0
23	N	5	0	3.24	8081.5	7117.2	6359.8	5793.4	5281.2	4835.5	4511.9	4254.7	3882.3	3662.8
25	Y	6	1	2.39	5960.9	5309.2	4816.1	4467.1	4152.5	3893.2	3706.0	3549.8	3335.6	3205.8
28	Y	6	2	1.42	3548.4	3091.6	2777.9	2525.3	2332.3	2141.9	1999.9	1914.1	1794.3	1702.1
41	Y	6	3	1.46	3627.7	3164.5	2837.3	2583.4	2382.8	2182.1	2037.4	1949.2	1822.4	1723.8
44	Y	5	1	2.03	5059.6	4371.2	3876.9	3504.7	3190.7	2920.4	2717.7	2578.8	2393.6	2258.3
48	Y	6	1	1.39	3460.9	2986.4	2653.9	2365.4	2163.6	1937.2	1772.6	1670.2	1527.5	1429.1
55	N	5	0	1.73	4307.6	3788.9	3451.1	3151.5	2934.7	2712.1	2529.6	2427.7	2300.8	2185.7
68	Y	6	1	0.96	2387.0	2095.3	1893.4	1742.8	1628.5	1522.9	1451.5	1398.5	1331.3	1286.5
73	Y	6	1	1.03	2556.9	2200.4	1953.5	1760.6	1613.6	1471.6	1367.8	1306.6	1220.1	1157.8
76	N	3	0	1.93	4807.7	4193.0	3762.8	3408.6	3137.2	2871.9	2666.8	2543.0	2372.5	2248.4
90	Y	5	2	2.10	5229.6	4563.7	4089.7	3723.0	3408.6	3129.7	2917.8	2773.5	2576.1	2428.0
95	Y	5	1	2.87	7153.8	6187.9	5536.7	4977.2	4618.0	4192.1	3907.9	3727.6	3493.9	3323.7
101	Y	5	2	2.09	5205.9	4496.6	3990.3	3601.5	3288.0	2987.8	2777.4	2638.1	2436.1	2301.3
102	N	3	0	2.06	5131.6	4534.5	4108.9	3774.4	3485.0	3238.4	3033.9	2887.6	2704.9	2569.1
110	N	5	0	1.89	4704.4	4160.8	3752.1	3416.9	3145.8	2880.7	2689.6	2547.2	2351.3	2226.0
2674	Y	6	2	1.80	4487.2	4020.8	3686.6	3402.7	3188.5	2955.3	2791.7	2674.5	2528.3	2417.1
2675	Y	4	1	1.69	4222.8	3612.3	3197.1	2821.6	2551.7	2281.0	2061.7	1952.5	1789.7	1660.6
2676	N	5	0	1.51	3759.1	3250.2	2893.9	2577.3	2365.1	2111.1	1935.2	1822.0	1669.9	1563.1
2678	N	5	0	1.54	3840.5	3326.1	2962.8	2677.5	2443.7	2242.3	2088.7	1983.6	1845.3	1747.6
2692	Y	6	1	1.28	3178.3	2800.2	2554.5	2345.3	2212.3	2049.9	1943.2	1873.2	1788.0	1716.2
2694	Y	5	3	1.28	3295.3	2782.6	2419.2	2130.0	1904.3	1689.8	1528.4	1429.6	1279.5	1181.8
2696	N	5	0	1.63	4066.5	3528.0	3148.5	2853.6	2618.2	2400.4	2243.6	2135.3	1989.2	1888.4
2698	Y	5	2	1.98	4929.6	4346.3	3944.4	3622.5	3389.8	3149.9	2979.4	2867.9	2717.1	2607.3
2700	Y	4	1	2.03	5061.8	4441.8	4010.5	3646.6	3377.5	3089.1	2886.1	2744.5	2562.8	2430.4
2706	Y	5	1	1.69	4203.5	3725.5	3409.7	3137.5	2952.3	2745.5	2589.6	2490.7	2363.3	2261.2

5.5 Figure Captions

Figure 1. An overlay of the objects in a small region (5000″ on a side) is shown to demonstrate the smoothing procedure. The region is from Field 681, and contains the two Abell clusters Abell 68 and Abell 73.

Figure 2. The test region, smoothed according to the procedure described in the text with $R_{\rm smooth} = 300''$. The black regions are those of high intensity, or equivalently, large galaxy density. Note how the two clusters easily stand out.

Figure 3. As in Figure 2, but with $R_{\rm smooth} = 600''$. This scale is a better match to these clusters' redshifts, hence they become even more apparent.

Figure 4. A histogram of the number of pixels having the values given on the abscissa for the smoothed images used to detect clusters, as in Plates 2 and 3. The value $R_{\rm smooth} = 420''$ was used. The curve with the highest peak is for the north field.

Figure 5. The profiles obtained when our profile procedure is followed on catalogues containing randomly distributed points. The random catalogues are constructed with the same surface density of objects as the actual galaxy catalogues, explaining the difference between the triangles and squares, which were computed for the north and south fields. The error bars shown are the standard deviations obtained from the ensemble of test positions at each smoothing radius.

Figure 6. The multiplicity function for clusters of galaxies is presented as a function of the overdensity parameter f computed using the local mean density of galaxies. The solid and dot-dashed lines are for the objectively selected clusters for the north and south field respectively. The squares and triangles indicate the Abell clusters present in the north and south fields. The multiplicity function of the objective clusters agree reasonably well, at least for rich (large value of f) clusters. Note the marked incompleteness of the Abell clusters compared to the objective ones at all values of f less than about 2.5.

Figure 7. Same as in Figure 6, but this time computing the overdensity parameter f with respect to the mean density observed in the south field.

Figure 8. We show the profiles for Abell clusters in the north field segregated into distance classes. Only the richness $R \geq 1$ (statistical sample) are included. The crosses are for distance class D = 4 clusters, the triangles for D = 5 and the squares for D = 6. The diamonds and solid lines indicate the profile derived from a catalogue of randomly distributed objects with the same overall mean density as the north field. Note the systematic difference in slopes between this reference profile and the clusters' profiles, due to the strong galaxy-cluster correlation function.

Figure 9. As in Figure 8, but for the south field.

Figure 10. Profiles measured for the north (solid lines) and south (dot-dashed lines) for Abell clusters of Distance classes $D = 5$ (triangles) and $D = 6$ (squares). The profiles shown are for the average of the clusters with the given distance classes and richness $R \geq 1$, from which we have subtracted the relevant background profile (the solid lines in Figures 8 and 9.) Note how different the profiles are, which indicate lack of homogeneitey of the Abell clusters within a given distance class.

Figure 11. We plot the measured values of f (explained in the text) for Abell clusters in the north field. The clusters are ordered on the abscissa by value of distance class D. Within each distance class, the objects with varying richness are represented by the size of the symbol, from $R = 0$ (very small symbol) to $R = 3$ (largest symbol). The correlation between the size of the symbol (within a given distance class) and its f value is very poor. The scatter is also very large, indicating a lack of reliability in the assignment of richness classes.

Figure 12. We plot the galaxy-cluster angular correlation function measured in the north (squares) and south (triangles). This was calculated for two sample of clusters. Very overdense clusters ($f \geq 1.9$) are shown as filled symbols, and less dense (but more numerous) clusters ($f \geq 1.5$) are shown with the open symbols. The lack of a clear scaling relation indicates that f is a very poor distance indicator. The shape of $w_{g-cl}(\theta)$ is remarkably similar to that predicted by the standard CDM model.

Figure 13. The predicted cluster-galaxy correlation function from a standard CDM simulation. Figure reproduced with the kind permission of the author.

5.6 Plate Captions

Plate 1. The frontiscpiece of this thesis, showing a greyscale representation of the ρ surfaces for both fields. The top panel is the north field, the lower panel the south. East is towards the bottom, and north towards the left. The fields were smoothed using $R_{\mathrm{smooth}} = 420''$, and galaxies in the magnitude range $17.5 \leq r \leq 18.5$ were used. The lookup table is the same for both fields.

Plate 2. Same as Plate 1, but showing only the north field. The orientation is as in Plate 1. We have superposed the position of the Abell clusters on the density map. Note how some clusters are on very weak density enhancement, and how many bright "hot spots" are not included in the Abell catalogue.

Plate 3. Same as Plate 2, but for the south field.

Figure 1.

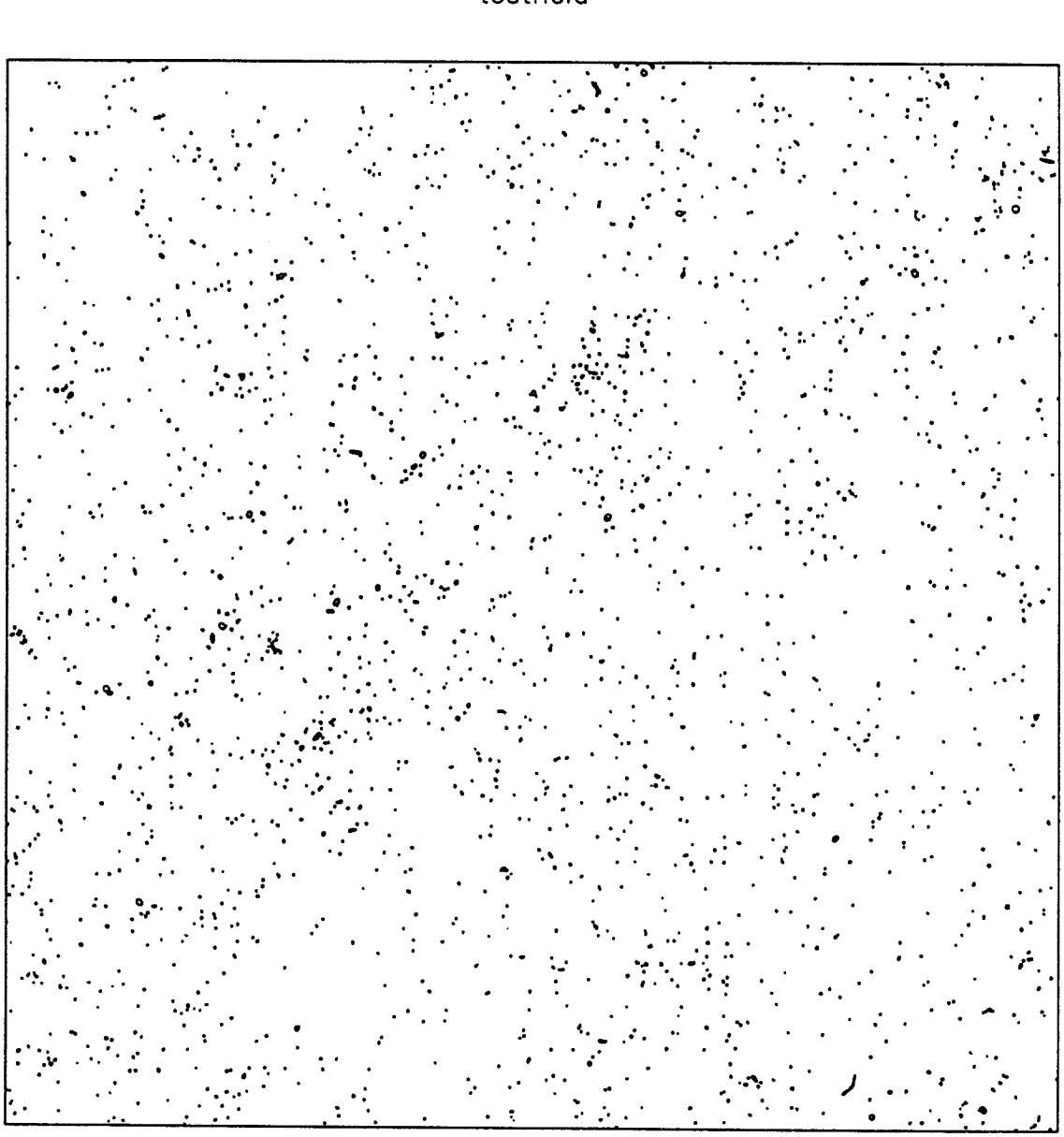


Figure 2.

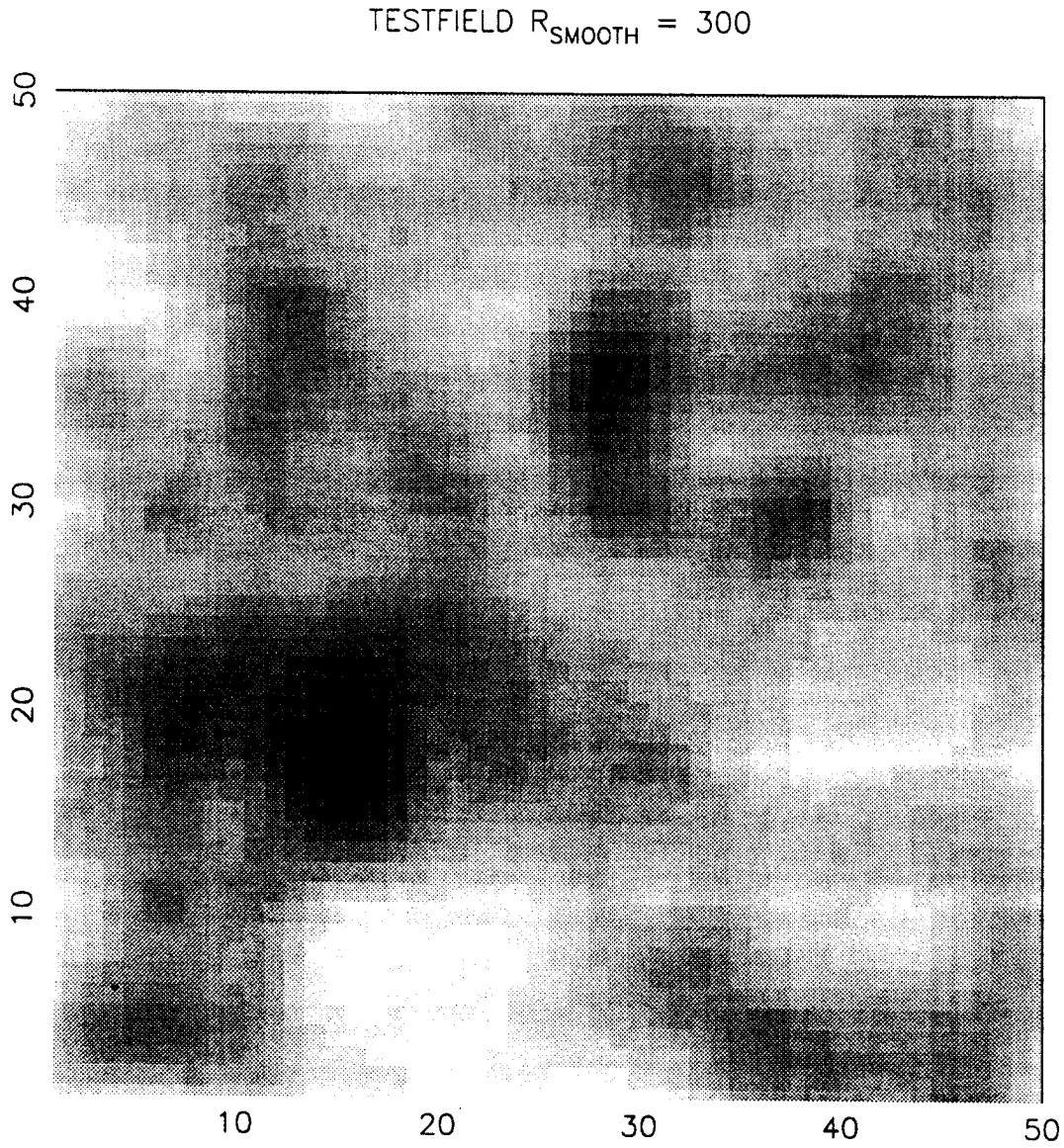


Figure 3.

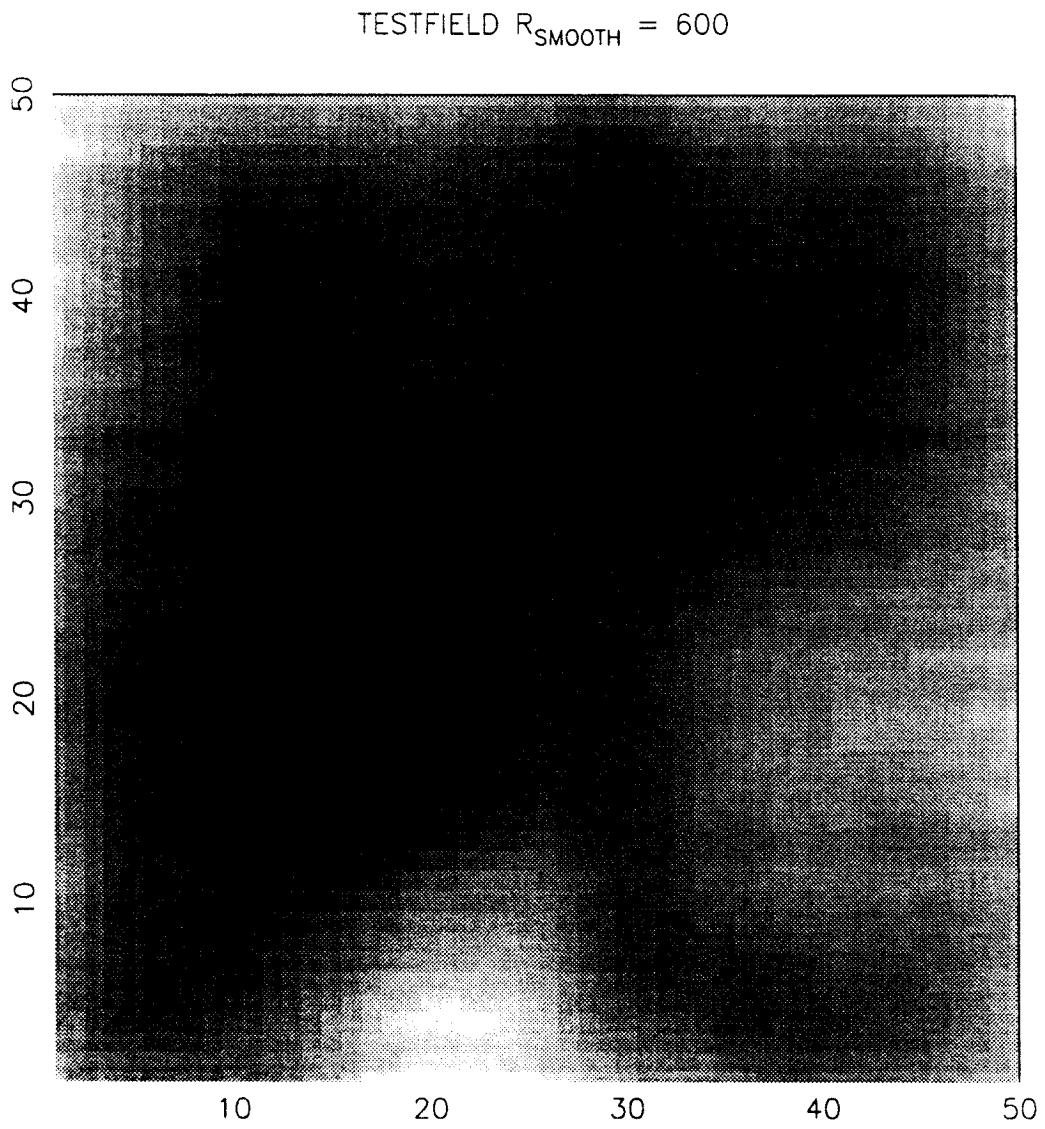


Figure 4.

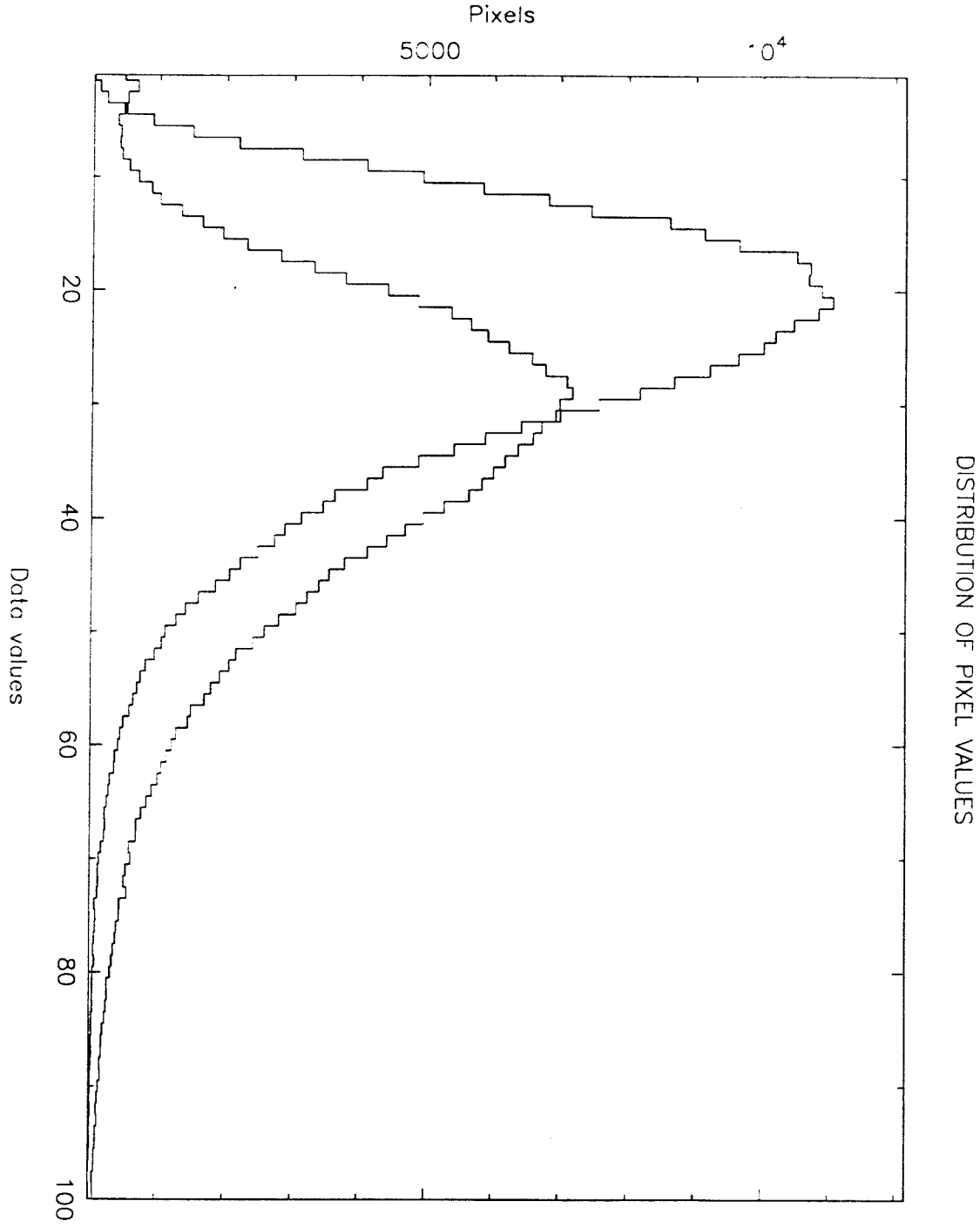

Pixels
5000
$\cdot 0^4$
20
40
60
80
100
Data values
DISTRIBUTION OF PIXEL VALUES

Figure 5.

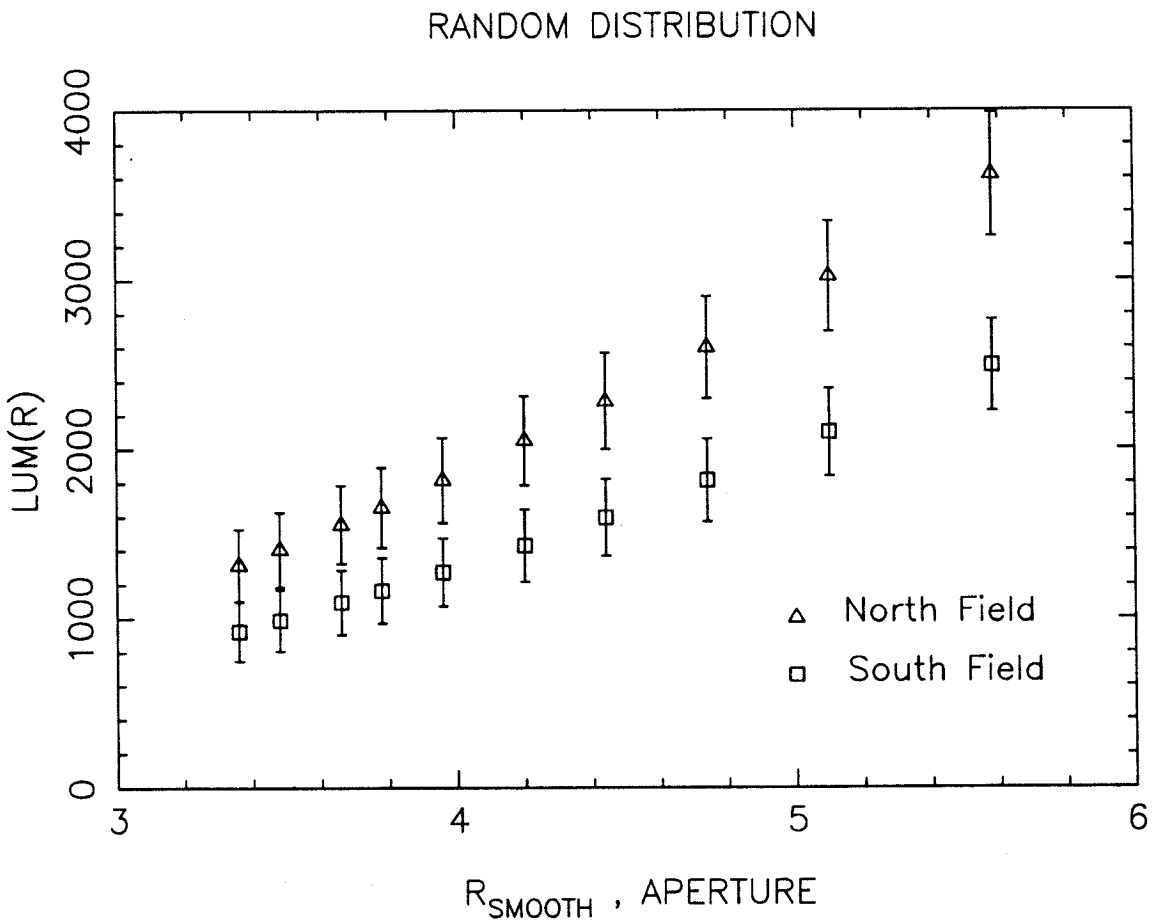

RANDOM DISTRIBUTION
LUM(R)
0
1000
2000
3000
4000
3
4
5
6
R_{SMOOTH} , APERTURE
North Field
South Field

Figure 6.

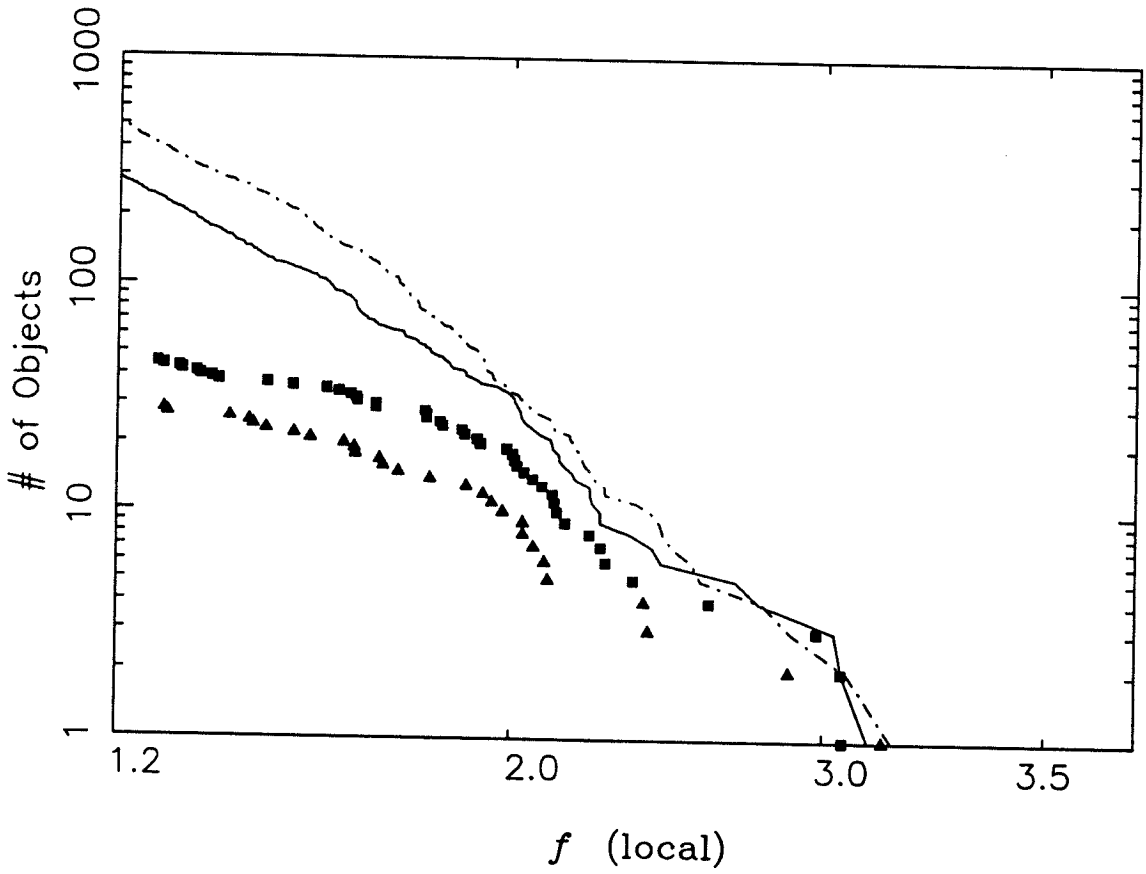


Figure 7.

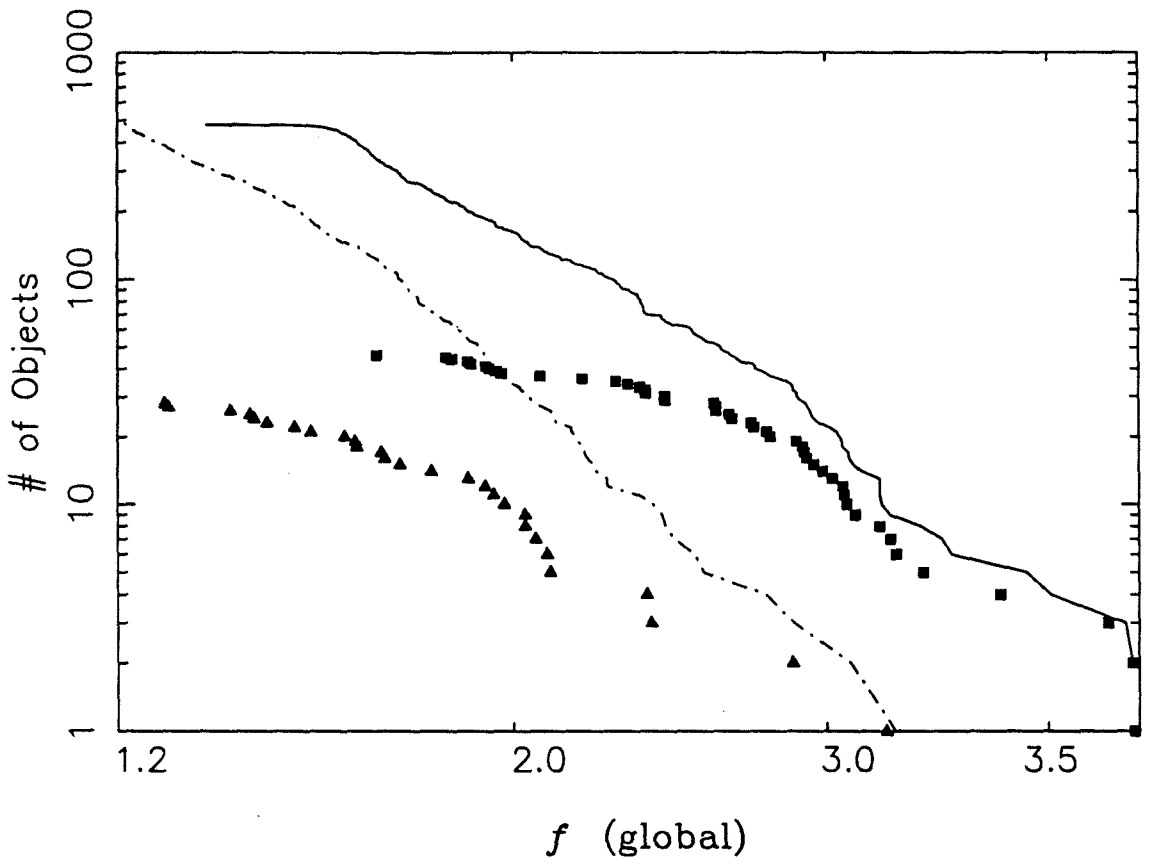


Figure 8.

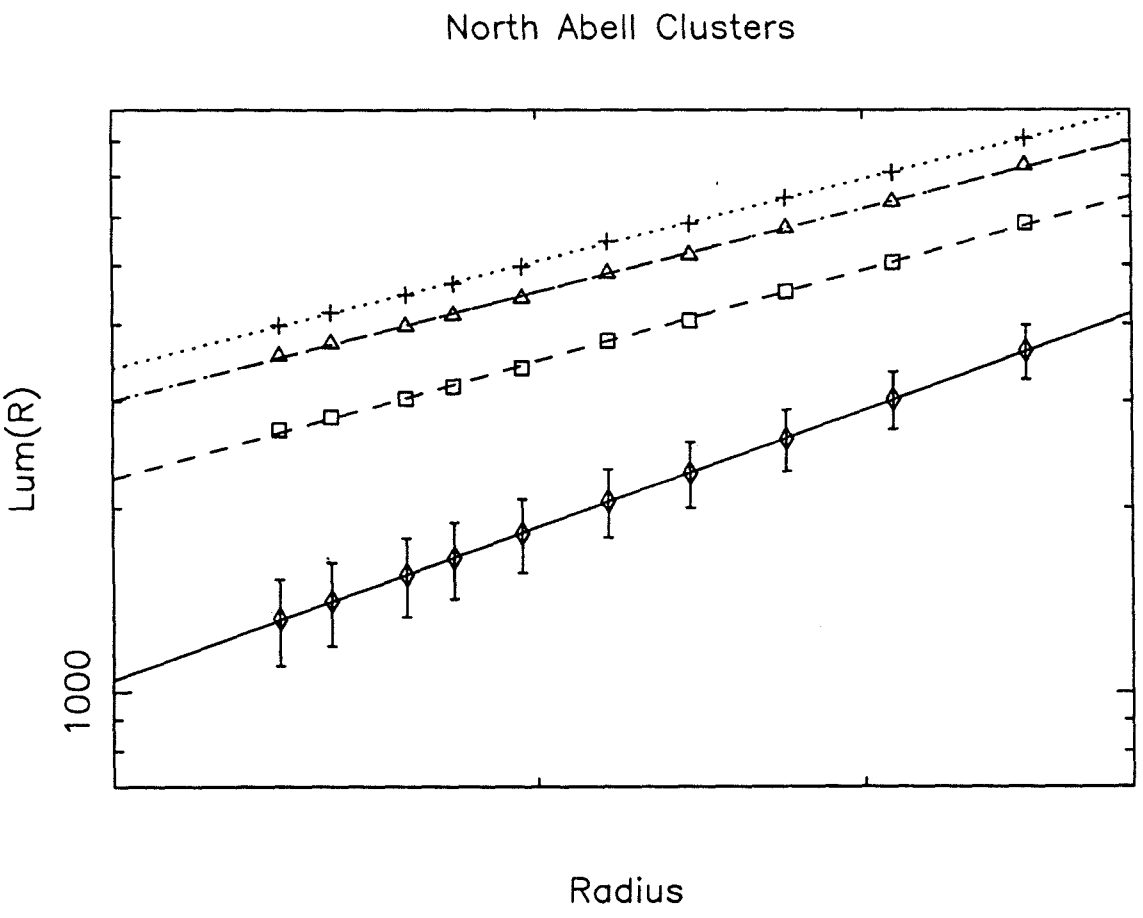


Figure 9.

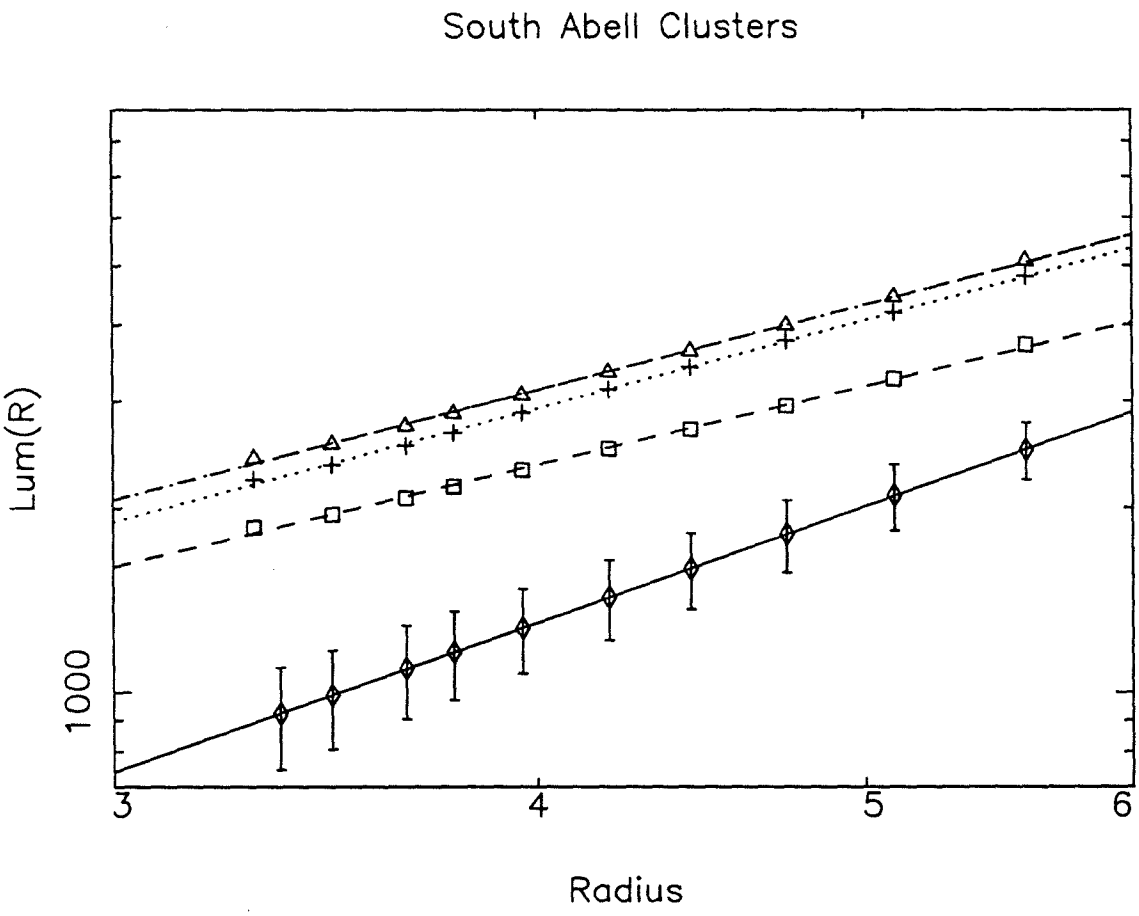


Figure 10.

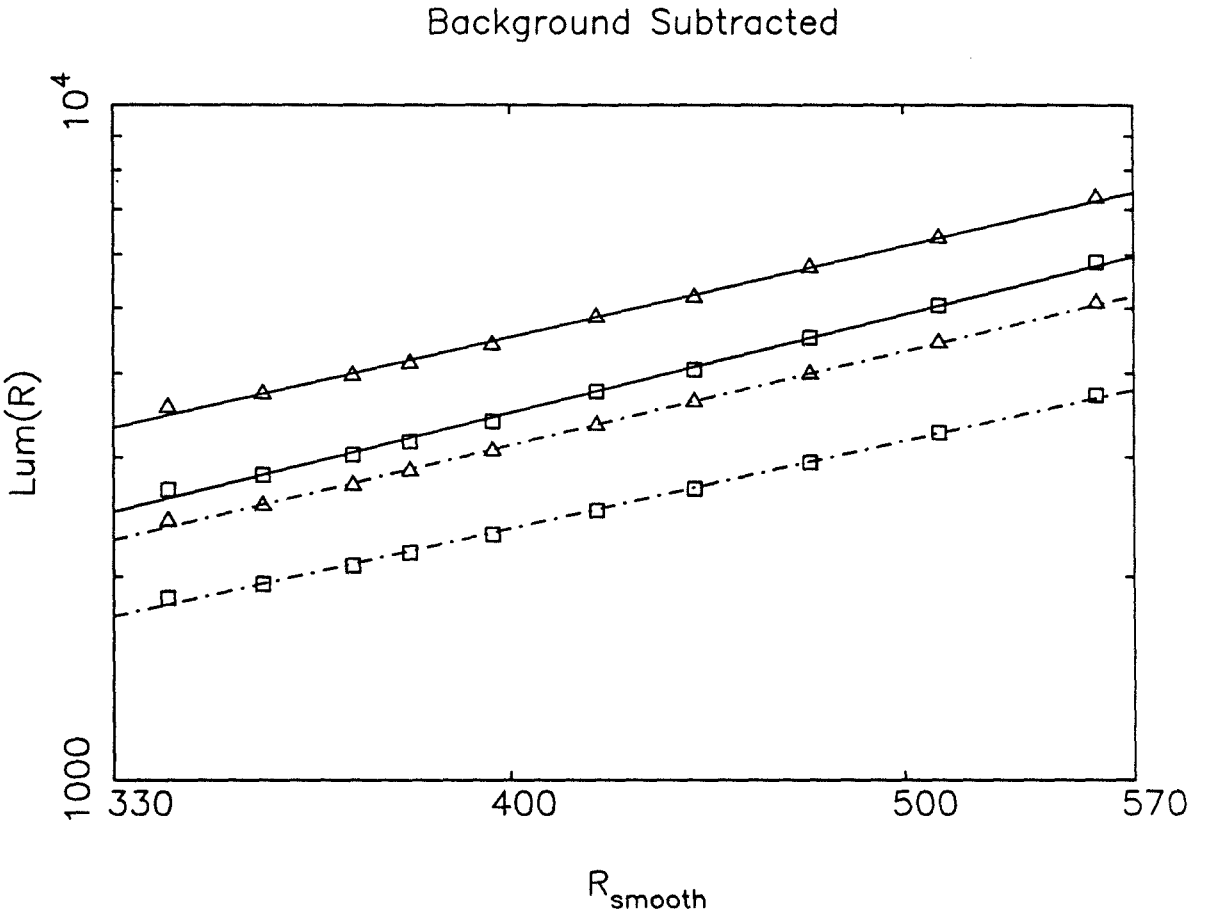

Background Subtracted
10^4
Lum(R)
1000
330
400
500
570
R_{smooth}

Figure 11.

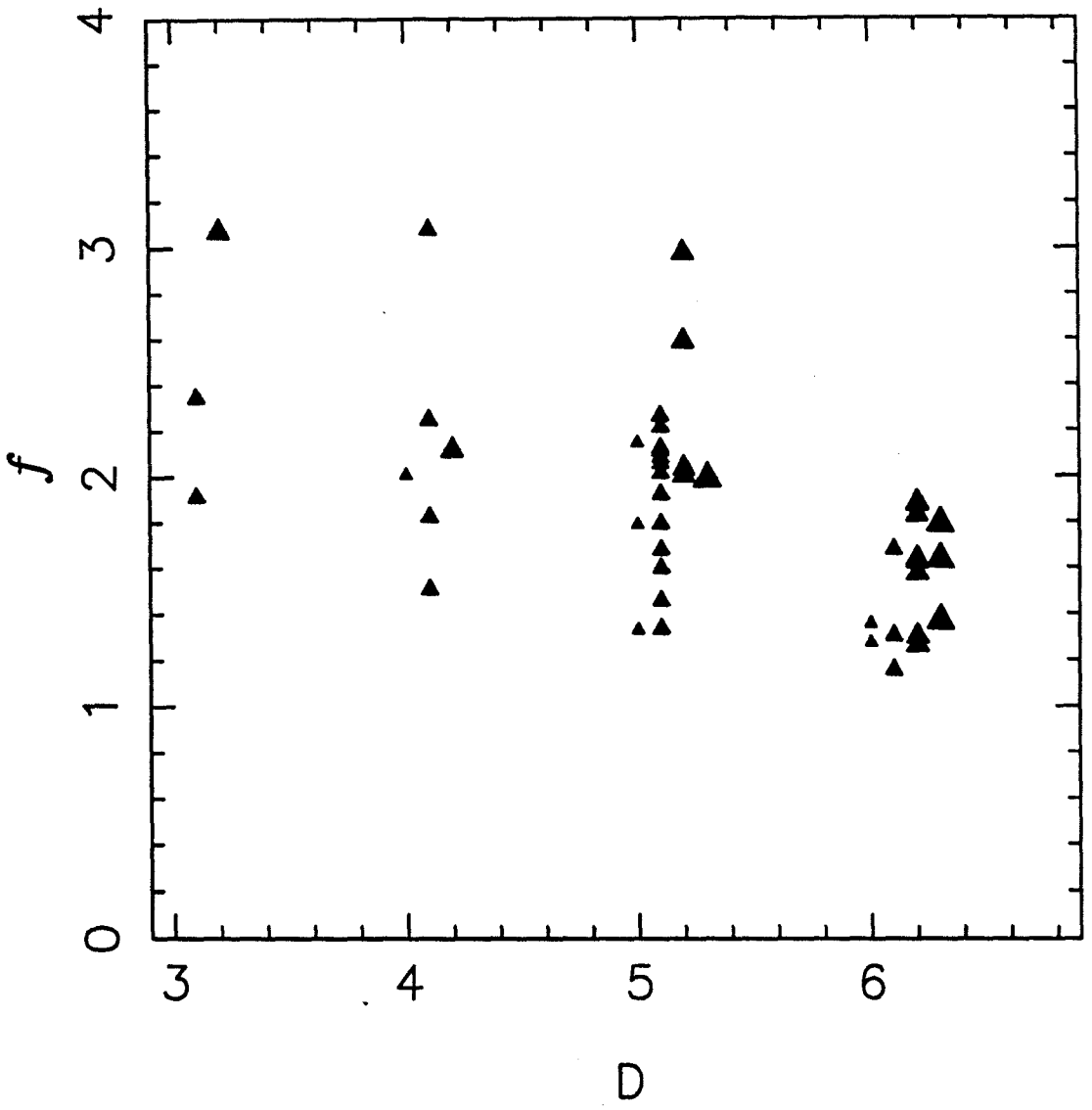


Figure 12.

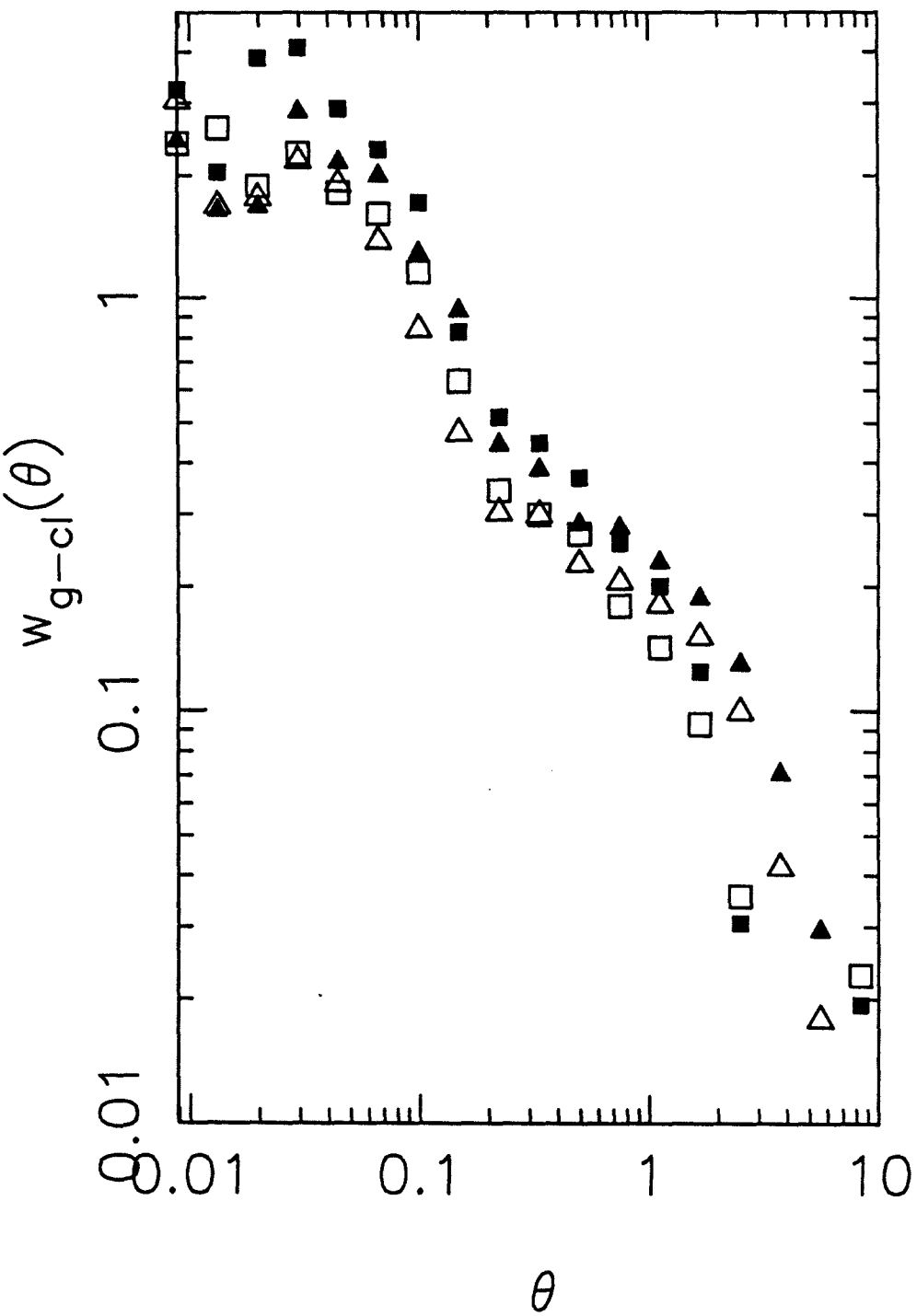
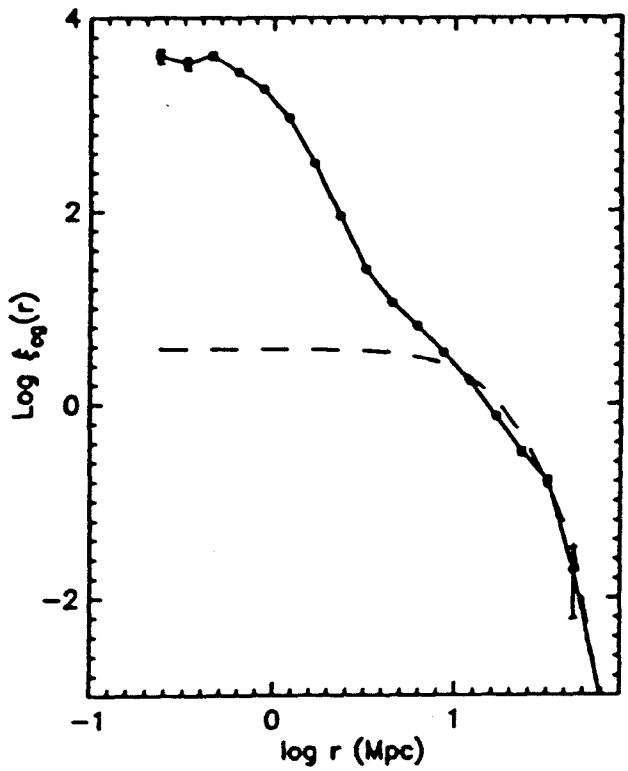


Figure 13.



5.7 References

Abell, G. O., 1958, *Ph.D. thesis.*

Abell, G. O., Corwin, H. G., Olowin, R., 1989, *Astrophys. J. Suppl.*, **70**, 1.

Bahcall, N.A. and Soneira, R.M., 1983, *Astrophys. J.*, **270**, 20.

Bardeen, J. M., Bond, J. R., Kaiser, N. and Szalay, A. S., 1986, *Astrophys. J.*, **304**, 15.

Dekel, A., Blumenthal, G. R., Primack, J. R., Olivier, S., 1989, *Astrophys. Letters*, **338**, 5.

Dodd, R. J., and MacGillivray, H. T., 1986, *Astron. J.*, **92**, 706.

Efstathiou, G., Sutherland, W. and Maddox, S. J., 1990, *Nature*, **348**, 705.

Lilje, P. B. and Efstathiou, G., 1988, *M.N.R.A.S.*, **231**, 635.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989a, *M.N.R.A.S.*, **242**, 43p.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989b, *M.N.R.A.S.*, **243**, 692.

McGill, C., and Couchman, H. M. P., 1990, *Astrophys. J.*, **364**, 426.

Postman, M., Geller, M. J. and Huchra, J. P., 1986, *Astron. J.*, **91**, 1267.

Rhee, G. 1989, *Ph.D. thesis.*

Shectman, S., 1985, *Astrophys. J. Suppl.*, **57**, 77.

Shortridge, K, 1984, *The FIGARO User's Guide, unpublished.*

Struble, M. F. and Rood, H. J., 1987, *Astrophys. J. Suppl.*, **63**, 543.

Sutherland, W., 1988, *M.N.R.A.S.*, **234**, 159.

West, M. J. and Van den Bergh, S.,1991, preprint.

6. Conclusion

6.1 What Has Been Achieved...

We have compiled a highly uniform catalogue of galaxies from COSMOS scans of POSS-II plates. The surveys covers an area of 386 square degrees, and covers 150,000 galaxies in the magnitude range $16.5 \leq r \leq 19.0$. The survey is divided into two regions, north and south galactic hemispheres.

The magnitude scale of each plate was brought to a common system using extensive CCD photometry, obviating the need for using the overlap regions in the calibration process. These overlaps were used to test the matching procedure, and it is found that systematic field effects may cause difficulty in the calibration of large surveys relying mostly on overlaps.

The procedure for separating stars and galaxies is completely objective. Visual tests indicate a 94% completeness of galaxies.

We have measured the angular correlation function of galaxies $w(\theta)$. The south field exhibits more power on large scales than the north field. We have performed several tests to determine the effect of possible systematic errors between the two fields and do not find any sufficiently large to account for the south field's excess power. Both fields are incompatible with the prediction of CDM. The north field is in good agreement with the results of the APM survey. Since the APM survey's CCD coverage is sparse, the agreement between the two surveys lends credibility to their result. The disagreement for the south field is not understood.

The number counts of galaxies in the two fields are substantially different. Several tests have been performed to make sure that the effect is not the result of a systematic error in zero point of the magnitude system or due to extinction or obscuration. Field 449, which contains the 6 Abell clusters which define the Corona Borealis supercluster has the highest density of galaxies of any plate; however, this supercluster itself cannot be responsible for the large difference in density between the north and south fields.

We put forward a toy model to determine the extent of the structure causing the over-density in the north (or under-density in the south). This model assumes a reasonable luminosity function and places the structure at a distance of $200h^{-1}$ Mpc. The natural scale for the structure is $50h^{-1}$ Mpc.

We have developed an objective algorithm for detecting clusters of galaxies from our galaxy catalogue. The algorithm is based on fixed overdensities at varying angular scales, and is well-suited to detecting clusters over a large range of redshifts. We present a catalogue of clusters based on this algorithm, using an overdensity threshold obtained from comparing the cluster's densities to each field's overall background density.

The results are very sensitive on the choice of background density. The multiplicity function of the objective clusters agree well when referenced to each field's background density, but disagree sharply when referenced to the overall mean density of the survey.

The objective catalogue is compared with that of Abell's. Poor agreement is found between the two catalogues. It is not possible to choose a set of parameters in our catalogue which reproduces Abell's catalogue well.

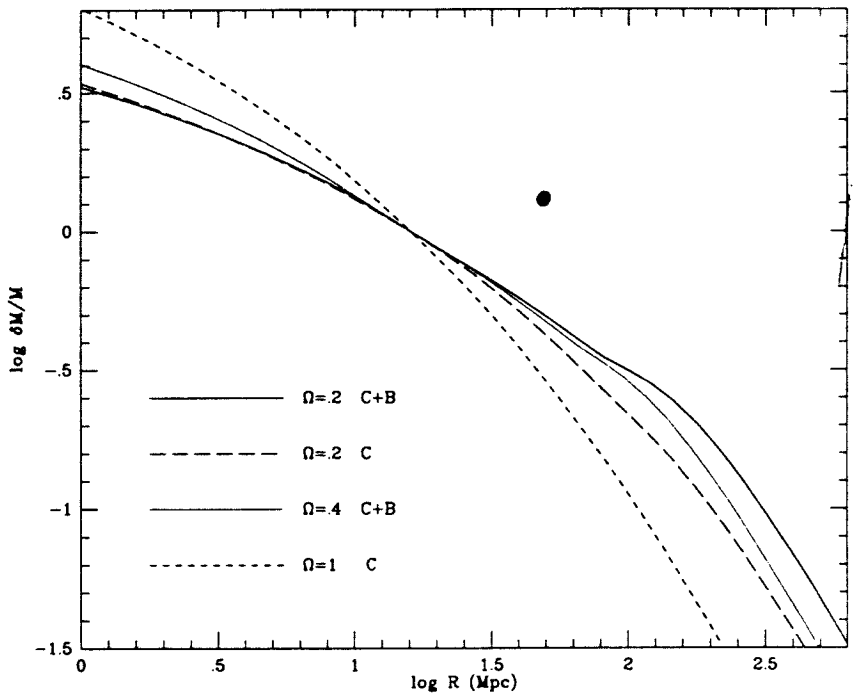
The measurement of the galaxy-cluster correlation function $w_{g-cl}(\theta)$ is in good qualitative agreement with predictions of CDM, but more detailed comparisons with theory must wait for the acquisition of redshifts for a large fraction of the clusters.

6.2 What It All Means...

The primary scientific results of this thesis are as follows. The power measured *via* w(θ) confirms the result of Maddox *et al.* (1989) and spells doom for standard CDM model.

Further, the large difference in number counts in the two fields allows us to place a point on the mass spectrum, as shown below:

Figure 1.



This figure is taken from Blumenthal *et al.* (1988), and shows the spectrum of density fluctuations in several CDM models. The "standard" model is represented by the short-dashed line. The authors adopt $h = 0.5$. We have added the solid dot at the location corresponding to $25h^{-1}$ Mpc (for $h = 0.5$) and at amplitude of δM/M = 1.3, which we obtain from our number counts assuming the scale in question is roughly $25h^{-1}$ Mpc (remember the total diameter of the structure indicated in Chapter 4 was $50h^{-1}$ Mpc), and we have assumed that the fluctuations in

the number of galaxies trace the fluctuations in the mass, which should be quite a reasonable assumption.

The figure speaks for itself.

It is entirely possible (in fact probable, if one takes the results of Maddox *et al.* 1989 at face value) that the difference in our two fields is an extremely rare event. The ordinate in the figure above is the *rms* value of the mass fluctuations. Clearly, it behooves us to obtain data in more fields to ascertain the significance of these observations.

And finally, the availability of objective catalogues of galaxies and clusters of galaxies will make it easier to compare theory with observations, by enabling the theorist to apply exactly the same selection effects to his model which face the observer in the real universe.

6.3 And What Remains to be Done

I would not be a true observational astronomer if I did not say that more and better data are needed. The cliché may be stale, yet it is nevertheless true.

The techniques for generating galaxy catalogues are now very good, and will in fact continue to improve as the amount of available computer power allows us to use more refined algorithms to perform image classification. What we need to do now is to improve the raw data. The first step, of course, would be to simply continue the present program over the whole sky. We suggest that all photographic surveys be accompanied by mandatory CCD surveys to calibrate the plates. This would enormously enhance the scientific value of these surveys.

The usable magnitude range of photographic surveys is small. This can be remedied in two ways. First, we might couple the present catalogues to ones derived from 4m prime focus plates to yield the desired dynamic range. The constraints on telescope time render this alternative unlikely, however.

The next generation of all sky transit surveys, planned to begin in the next few years, will go a long way towards solving these problems. The great stability of CCD detectors, and their enhanced dynamic range, will yield deeper and more uniform surveys. I am jubilant at the thought of Terabytes of raw data needing to be crunched...

A program is already under way to obtain redshifts for as many of the objective clusters as possible. These will yield insight into the physical properties of the objects found by our algorithm (in particular, we would like to know which fraction of the objects are just chance projection of galaxies and are not physically associated) and may allow us to determine the cluster-cluster correlation function once and for all.

The galaxy catalogue may serve as the basis for a complete redshift survey (a deeper "slice of the universe"), or be used to investigate internal cluster dynamics through the use of multi-object spectrographs, such as the Norris spectrograph.

Soon, we will have a cluster catalogue from ROSAT. This catalogue should be essentially free from contamination effects due to projection. Comparison of such a catalogue with ours may yield insight as to which objects are real, and how well one can define clusters from 2D surveys.

It is becoming apparent that we have taken the two point correlation function to its limits. This statistic was made popular by Groth and Peebles' (1977) analysis of the Lick counts, and has been a standard tool for the cosmologist ever since. Its main virtue is that it is easy to compute, but advances in computer technology have rendered this reason obsolete. Obviously, describing the entire 2-D or 3-D distribution of a set of points using a slope and an amplitude is a gross simplification. I think we will see emerge in this decade better statistical tests for comparing observations to theoretical models. Work has already begun in applying the genus statistic of Park and Gott (1990) to our catalogue. We are currently also applying void probability statistics, and counts in cells (Quashnock and Picard 1991), which contain information on all order correlation terms.

The results contained in this thesis and elsewhere constitute a challenge to the observers, to keep mapping the universe in search of perhaps yet larger structures, or to determine on what scales the universe finally *does* become homogeneous. They are also a challenge to the theorists, who must now revise their models to account for these observations.

This will indeed be a busy decade for cosmologists...

6.4 References

Blumenthal, G. R., Dekel, A. and Primack, J. R., 1988, *Astrophys. J.*, **326**, 539.

Gott, J. R. III, Park, C. B., Juszkiewicz, R., Bies, W. E., Bennet, D. P., Bouchet, F. R., and Stebbins, A., 1990, *Astrophys. J.*, **352**, 1.

Groth, E. J. and Peebles, P. J. E., 1977, *Astrophys. J.*, **217**, 385.

Maddox, S. J., Efstathiou, G., Sutherland, W. J. and Loveday, J., 1989, *M.N.R.A.S.*, **242**, 43p.

Quashnock, J. M., and Picard, A., 1991, in preparation.