Mon. Not. R. Astron. Soc. **329,** 76–86 (2002)

Variability in active galactic nuclei: confrontation of models with observations

M. R. S. Hawkins★
University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ

Accepted 2001 August 28. Received 2001 April 18

ABSTRACT

The variability of active galactic nuclei (AGN) has long held the promise of shedding light on their detailed structure, and possibly other astrophysical phenomena. Different emission mechanisms lead to different patterns of variability in flux, which are in principle easily distinguishable. Recent predictions for the expected spectrum of variations for various models are now in such a form that they can be compared with the observed statistical properties of AGN light curves from large-scale monitoring programmes. In this paper, we use the results of a long-term monitoring programme of a large sample of quasars and Seyfert galaxies, as well as individual light curves from the literature, to distinguish between the various model predictions. The results favour a model based on accretion disc instability over the starburst model, where the variation comes from a succession of supernova bursts, but it also appears that much of the observed variation in quasars is due to gravitational microlensing.

Key words: galaxies: active – quasars: general – galaxies: Seyfert – cosmology: miscellaneous.

1 INTRODUCTION

Forty years after the discovery of quasars there is still much uncertainty about their structure. There is, arguably, broad agreement about the basic nature and overall arrangement of the various components of a 'unified model' for active galactic nuclei (AGN) (Antonucci 1993), but the details have proved very hard to tie down. In particular, the nature of the central engine and the radiative transfer processes are the subjects of much debate. Part of the problem is the paucity of observations that can effectively distinguish one model from another. Two types of constraint that are of particular interest are the spectral energy distribution and the observed variations in flux. Although one can deduce much about AGN structure from properties such as the optical/X-ray flux ratio (George & Fabian 1991) or the so-called 'big blue bump' (Gondhalekar et al. 1994), such measurements are not sufficient to distinguish between competing models, let alone refine model parameters.

Variations in flux were detected in quasars shortly after their discovery, and right from the start have played a pivotal role in constraining quasar morphology. The early detection of light fluctuations on a time-scale of months provided a fundamental constraint on the size of the emitting region which has underlain all subsequent efforts to model the structure of quasars. In order to obtain a clearer picture of the emitting regions in AGN and the associated emission mechanisms, a number of monitoring programmes have been carried out to measure the spectrum of variations in different wavebands.

At present there are three basic models for explaining the observed AGN variability. The first involves instabilities in the accretion disc, the central engine powering the energy output (Rees 1984). The second postulates that AGN are powered by multiple supernova explosions or starbursts which result in stochastic variations in brightness (Aretxaga & Terlevich 1994). In the third approach, the observed variations are not intrinsic to the AGN at all, but the result of gravitational microlensing by small compact bodies or massive compact halo objects (MACHOs) along the line of sight (Hawkins 1993). In fact, it seems quite likely that two or even all three of these processes are present at some level, and so the task of the observer is to disentangle them all in order to draw useful astrophysical conclusions. There is also the possibility that different mechanisms dominate in different luminosity regimes, and so in the analysis in this paper we shall divide AGN into two categories, quasars with $M_B < -23$ and Seyfert galaxies with $M_B > -23$.

Although much effort has been put into monitoring quasars (Trevese et al. 1994; Hook et al. 1994; Cristiani et al. 1996; Hawkins 1996), both individually and in samples of various sizes and different selection criteria, it seems fair to say that little light has so far been shed on the nature of quasars using these methods. There appear to be two main reasons for this. First, despite the efforts that have been put into quasar monitoring programmes, the data have on the whole proved inadequate for measuring useful

★E-mail: mrsh@roe.ac.uk

parameters to characterize the variability. This is partly because samples of quasars have been too small, and partly because the run of data has been too short and inhomogeneous. As a consequence it has not proved possible to define unambiguously the fundamental properties of the flux variations of the quasars.

The second reason that quasar monitoring programmes have not led to more progress in the understanding of AGN concerns the lack of firm predictions for variability from the various competing AGN models. A big step forward in this area has recently been made with the publication (Kawaguchi et al. 1998) of a detailed model for AGN variability from accretion disc instability. In that paper the authors make detailed statistical predictions for the spectrum of fluctuations for their model and also for the starburst model. These predictions are presented in a form that enables meaningful comparison with observations of AGN variability.

It is the purpose of the present paper to use the best available observational data of the variations of AGN to distinguish between the various models of variability. To this end we describe a large-scale monitoring programme of a sample containing some 600 quasars with regularly sampled light curves covering 24 yr. We also use extensive monitoring data for Seyfert galaxies taken from the recent literature. Predictions of the models are published in the form of structure functions, and we analyse the observational data in the same way to provide quantitative comparisons.

2 THE STRUCTURE FUNCTION

In order to extract information from observations of AGN variability, it is necessary to find ways of quantitatively characterizing the nature of the variations. There is a vast literature on time-series analysis, much of it concerned with extracting information from incomplete or inhomogeneous data sets. There are several functions that have been used for analysing AGN variations, including the structure function, autocorrelation function and Fourier power spectrum. Although all of these functions contain very similar information, in practice they each have advantages in particular situations. For example, although, for an infinite and complete run of data, the autocorrelation function and power spectrum are essentially equivalent to each other, for finite runs of data there are significant differences. For long runs of evenly spaced data, the power spectrum is to be preferred, as it is on the whole easier to interpret and understand the errors. For short or inhomogeneous data sets, the autocorrelation function provides a more stable measurement, but as the individual points are not independent of each other, there can be difficulties with interpretation. The structure function is very similar to the autocorrelation function and has been widely used in the analysis of quasar light curves (Trevese et al. 1994; Hook et al. 1994; Cristiani et al. 1996; Aretxaga, Cid Fernandes & Terlevich 1997) and microlensing statistics (Wyithe & Turner 2001). The function of choice will depend upon a number of factors, not least of which is the form in which model predictions have been made.

The present paper has been largely prompted by the publication (Kawaguchi et al. 1998) of quantitative predictions for the statistics of AGN variability, and these are presented in the form of structure functions. Accordingly, we shall proceed with the analysis of the observations in the same way. The structure function S may be defined by

$$S(\tau) = \left\{ \frac{1}{N(\tau)} \sum_{i<j} [m(t_j) - m(t_i)]^2 \right\}^{1/2},$$

where $m(t_i)$ is the magnitude measured at epoch t_i, and the sum runs over the $N(\tau)$ epochs for which $t_j - t_i = \tau$. The interpretation of the structure function in the sense of identifying specific characteristics of the variation is not usually feasible. However, particular models of variability can be shown to produce structure functions with measurable parametric forms. Although in some cases efforts have been made to predict the shape of the structure function (Cid Fernandes, Terlevich & Aretxaga 1997), it is perhaps more useful to generate structure functions from simulated data (Kawaguchi et al. 1998), and this is the approach underlying the present paper. In addition to the standard structure function defined above, we shall for the purpose of measuring asymmetries also make use of two modified structure functions S_+ and S_-. These are defined as for S, except that for S_+ the integration only includes pairs of magnitudes for which the flux becomes brighter, and for S_- for which it becomes fainter.

3 MODELS OF AGN VARIABILITY

The mechanism behind the variability of AGN has been the subject of much debate since their first discovery. At present there are three broad approaches to explaining the observed variations. The most favoured model has invoked instabilities in the accretion disc surrounding the central black hole, but a contending idea is that the variations are caused by some intermittent sequence of discrete events such as supernova bursts. A third possibility is that we are seeing the effects of gravitational microlensing by a population of small compact bodies along the line of sight. Each of these models has good arguments in its favour, but also fails to explain some aspects of AGN variability, and it seems likely that all three processes contribute at some level to the observed light curves, with perhaps one mechanism dominating in any particular regime. Until recently, there has been little attempt in the literature to make model predictions that can easily be tested against observations, but that is now changing. In the remainder of this section we review the current position for the three basic models mentioned above.

3.1 Disc instability model

The description of the accretion disc model as proposed by Rees (1984) contains much discussion about time-scales of variability, but makes no predictions about the spectrum of variations. Since then there has been much work on variability in accretion discs (Wallinder, Kato & Abramowicz 1992), but the relation between the various types of instability and changes in emitted flux have proved hard to tie down and turn into specific testable predictions for an observed spectrum of variations. In fact, to the extent that predictions are possible, they would appear to be at variance with much of the observed variability.

A major step forward has been made in a recent paper by Kawaguchi et al. (1998), who develop the cellular-automaton model for disc instability of Mineshige, Ouchi & Nishimori (1994) to produce quantitative testable predictions for the shape of the structure function for the resulting photometric variations of the disc. The basic idea of the cellular-automaton model is that, as matter flows through an accretion disc, it causes instabilities that produce a spectrum of avalanches of different sizes, which in turn manifest themselves as variations in emitted flux. The time-scale of the light variation corresponds to the accretion time-scale τ_{acc}, which is defined (Kawaguchi et al. 1998) as

$$\tau_{\mathrm{acc}} = 160\left(\frac{r}{10^2 r_{\mathrm{g}}}\right)^{3/2}\left(\frac{M}{10^9\,\mathrm{M}_{\odot}}\right)\,\mathrm{d},$$

where r_{g} is the Schwarzschild radius. For plausible values of the black hole mass M this implies a time-scale of days or less for emission from the innermost stable orbit. However, observations of quasar light curves imply time-scales of at least several years (Hook et al. 1994; Hawkins 1996; Cristiani et al. 1996), which for the cellular-automaton model would mean emission from a characteristic radial distance of $\sim 1000 r_{\mathrm{g}}$. In the notation of Kawaguchi et al. (1998), this would imply $r_{\mathrm{in}}/r_{\mathrm{out}} \approx 1000$. The authors use Monte Carlo simulations of the model for several parameter sets to produce simulated light curves from which they can evaluate structure functions. The structure functions have the form of a power law that flattens at the time-scale τ_{acc}. It may be seen from their table 2 that the logarithmic slope of their structure function is not sensitive to the ratio $r_{\mathrm{in}}/r_{\mathrm{out}}$ or to their other input parameter m'/m, the ratio of diffusion mass to inflow mass, but lies close to 0.44 ± 0.03. This is a stable figure that can thus be compared with observations.

3.2 Starburst model

The idea that AGN variability might be caused by a series of discrete outbursts, such as supernova explosions superposed as a Poisson process, has a long history. It was however recognized early on that there was a fundamental problem with the idea, which had come to be known as the ‘Christmas tree’ model (Pica & Smith 1983). The model predicts a relation between the luminosity and amplitude of quasars which is not observed, and so was ruled out early on as a serious contender for explaining their structure and variability. Despite this the model has some attractive features, especially for explaining low-luminosity AGN such as Seyfert galaxies where the nuclear luminosity could plausibly be explained by a few supernovae per year, which could at the same time account for the observed amplitude of variability. For quasars of even moderate luminosity the required supernova rate of around one per day is such that the amplitude of variation would be far too small to be consistent with observations. Notwithstanding this, Terlevich et al. (1992) made a good case that the spectroscopic properties of AGN could be explained as originating from supernovae with their ‘starburst’ model. The case that the variability of quasars can be explained (Aretxaga & Terlevich 1994) is not convincing, but it cannot be dismissed as a possible dominating feature of Seyfert galaxies.

One positive feature of the starburst model is the feasibility of predicting the spectrum of variations that will be observed. Kawaguchi et al. (1998), in addition to the accretion disc model, also produced model light curves and structure functions for starbursts. They used the formalism of Aretxaga & Terlevich (1994) to construct a model and perform Monte Carlo simulations. Their fig. 3 shows model light curves for a quasar of moderate luminosity which well illustrates the fundamental problem with the starburst model as applied to quasars. The amplitude of variation is about 0.2 mag, which is far less than would be expected for such an object (Hawkins 2000).

The structure functions of the model starburst light curves have a power-law form with a flattening at around 100 d, depending on model parameters. The power-law section of the structure function has a logarithmic slope of about 0.83 ± 0.08, much larger than for the accretion disc, and a potential discriminant between the two models.

3.3 Microlensing model

The third broad approach to explaining observed AGN variability invokes gravitational microlensing. Unlike the two models already described, no claim can be made that microlensing explains all variations seen in AGN. The hypothesis is that for quasars the observed long-term variations (on a time-scale of several years) are dominated by the effects of microlensing. Here, the main interest centres on the microlensing bodies, although there is the potential to learn about the structure of the central region of quasars from the microlensing process.

The idea (Hawkins 1993, 1996) is that the dark matter is made up of a large population of planetary-mass compact bodies or MACHOs, which behave as cold dark matter and randomly cross the line of sight to any particular quasar. The optical depth to microlensing is such that on average a quasar’s light is lensed by several bodies at any one time, resulting in a distribution of caustics that causes a complicated amplification pattern on a time-scale of a few years. Low-redshift Seyfert galaxies are too nearby for any significant probability of microlensing, and so any variation seen in them must be intrinsic. Microlensing is well known to occur in multiply lensed quasar systems, where some variations are only seen in one image, and on long time-scales appear to dominate over the fluctuations seen in all images, which must be intrinsic to the quasar (Pelt et al. 1998).

Over the last 20 yr or so a number of groups have published computer simulations of the light curves produced by microlensing (Kayser, Refsdal & Stabell 1986; Schneider & Weiss 1987; Lewis et al. 1993). Fig. 1 shows data from fig. 2(a) of Lewis et al. (1993) of magnitude versus Einstein radius R_{E}, sampled to enable comparison with observed light curves. Structure functions were calculated for these data, and also for the light curves for model b from Schneider & Weiss (1987). The two simulations are for point sources, but assume different surface mass density κ_*, which is essentially equivalent to the optical depth to microlensing (Schneider & Weiss 1987). The two structure functions are plotted in the left (Schneider & Weiss 1987) and right (Lewis et al. 1993) panels of Fig. 2. The structure functions have logarithmic slopes of 0.28 and 0.23 respectively over the linear part, and there appears to be more power for greater optical depth.

In microlensing simulations, the source size has a very noticeable effect on the appearance of the light curves, which have a smoother, more rounded structure as it becomes larger. Fig. 3 shows the structure function for data from fig. 2(b) of Lewis et al. (1993). These data come from the same distribution of lenses as for their fig. 2(a), but with a source size of $0.2R_{\mathrm{E}}$. The logarithmic slope of the linear part of the structure function in Fig. 3 is 0.31. This is slightly more than for a point source, as is to be expected from the loss of high-frequency components attributable to a resolved source, although there is only a small difference in amplitude at long time-scales.

3.4 Comparison of model predictions

It is interesting that the structure functions for all three models discussed above have a similar morphology. There is a power law or logarithmically linear rise from the shortest time-scales to an eventual long time-scale break or turnover. In each model the power or amplitude of the structure function is somewhat

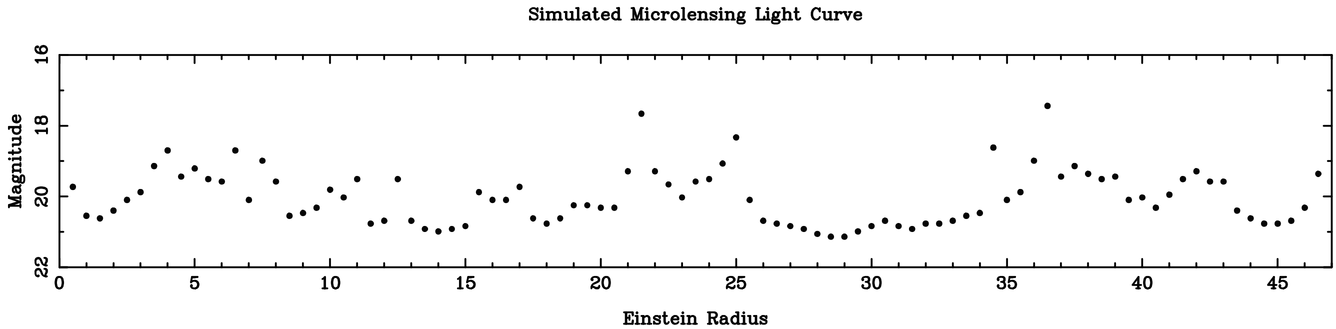

Figure 1. Simulated microlensing light curves from fig. 2(a) of Lewis et al. (1993).

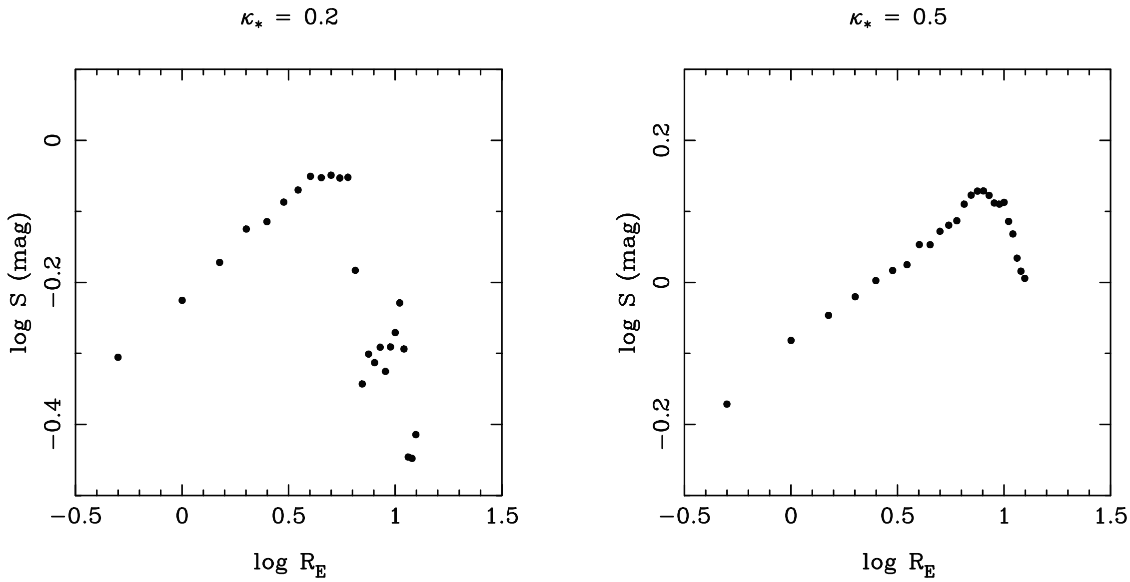

Figure 2. Structure functions for simulated microlensing light curves for model b of Schneider & Weiss (1987) (left-hand panel) and from fig. 2(a) of Lewis et al. (1993) (right-hand panel).

dependent on the choice of model parameters, but broadly speaking appears to be less for disc instability than for microlensing, and very much less for starburst. The time-scales are difficult to compare, as it is only the starburst model that makes predictions for a specific time-scale, linked to the supernova cooling time-scale, which is well known, a characteristic value being 280 d (Kawaguchi et al. 1998). The accretion disc model is defined in terms of a time-step, which is left as a free parameter. Its value is related to the accretion time-scale, which in turn depends on the size of the emitting region of the accretion disc. The time-scale of microlensing models depends on the Einstein radius of the lenses and their mean transverse velocity. Although the second parameter can be estimated with reasonable confidence, R_E is directly related to the mass of the lenses, which in general is completely unknown.

It is remarkable that the logarithmic slopes of the linear part of the structure functions show only small dispersions for different choices of parameter within each of the three models described above, but that the means are well separated. The slopes are 0.83 ± 0.08, 0.44 ± 0.03 and 0.25 ± 0.03 for the starburst, accretion disc and microlensing models respectively, which makes for a good opportunity for distinguishing between the models by comparison with observations.

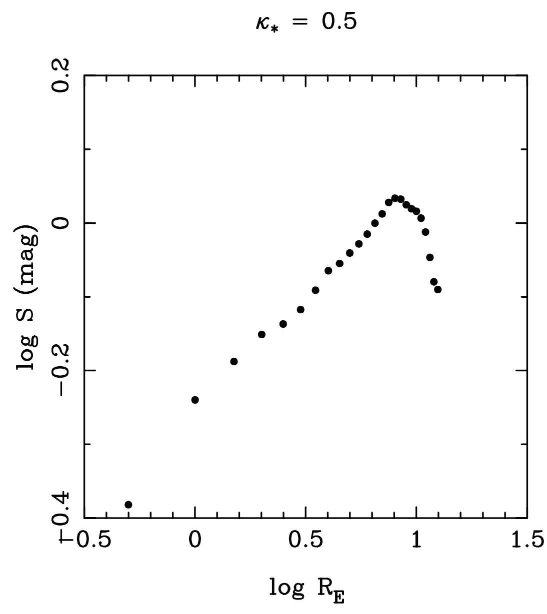

Figure 3. Structure function for the simulated microlensing light curve from fig. 2(b) of Lewis et al. (1993).

4 OBSERVED STRUCTURE FUNCTIONS

4.1 Quasars

Over the past ten years or so there have been several major projects to monitor the light variations of quasars on a time-scale of ten years or more (Trevese et al. 1994; Hook et al. 1994; Hawkins 1996; Cristiani et al. 1996). The data for these monitoring programmes have been presented in the form of structure functions or autocorrelation functions, and used to look for correlations of redshift and luminosity with various parameters related to amplitude and time-scale. It seems fair to say that the search for a correlation with the amplitude or power of the structure function has been a lot more successful than questions involving time-scale of variation.

The problem with the existing data is that the structure functions are not sufficiently well defined to compare with or distinguish between model predictions. There are a number of reasons for this. Perhaps the most serious difficulty is the smallness of the time-span over which the quasars have been monitored, which means that features in the structure functions such as breaks or slopes are not defined with sufficient accuracy. A related problem concerns lack of regularity in the observations. Unevenly sampled light curves are notoriously difficult to analyse, resulting in various types of aliasing and artefacts that can be very misleading in their interpretation. A different problem concerns sample sizes. The difficulty of monitoring a large number of quasars individually on a regular basis is such that most monitoring projects have been based on wide-field surveys, especially Schmidt telescopes. Even so most samples of quasars have been limited by available redshifts to around 300 members or less. A final limitation has concerned the analysis procedures used, which have been tailored to the unevenly spaced data that are hitherto all that have been available.

In this paper we present data that go a long way to rectifying the problems mentioned above. The light curves are part of a monitoring programme that has been in progress since 1975, based on repeated measures of a single 1.2-m UK Schmidt Telescope field (ESO/SERC Field 287) centred on $21^{\mathrm{h}}28^{\mathrm{m}}\ -45°$. The field has been observed several times every year since 1977 in the B_{J} passband (Kodak IIIa-J emulsion with a Schott GG395 filter), as well as on shorter time-scales of months, weeks and days, and also in other passbands, notably R and U. In this paper we concentrate on the measures in B_{J}, which form an unbroken sequence of 24 yearly measures from 1977 to 2000. At least one deep exposure was obtained every year during this period, and in most years the measures are based on four or more exposures. The plates were measured on the COSMOS or SuperCOSMOS automated measuring machines at the University of Edinburgh to produce yearly magnitudes for the $\sim 200\,000$ objects in the central 19 deg^2 of Field 287. The photometric error on a measurement from a single plate was approximately 0.08 mag, and for most years when four plates were available this reduced to 0.04 mag. More basic details of the survey, including a discussion of measurement errors, have already been published (Hawkins 1996, 2000), and here we concentrate on updating aspects of the survey relevant to the present paper.

The quasars in the field have been detected using a number of techniques including ultraviolet excess, red drop-out, variability, objective prism searches and radio surveys. There are estimated to be about 1500 quasars in the field to a magnitude limit of $B_{\mathrm{J}} = 22$, of which 610 have been confirmed with redshifts. Of these, several complete samples have been compiled according to various well-defined criteria (Hawkins 2000).

The light curves show a variety of features, especially on a time-scale of several years or more. Fig. 4 illustrates some examples. For most years the plotted magnitudes are the mean of four measures, and the error bars are based on the average photometric errors and the number of plates available in any particular year, which in most cases was four. The light curves show no obvious asymmetries in time, or easily definable morphological characteristics. Perhaps the one thing that can be said is that there appears to be more power on longer time-scales, and that more time is needed to be sure that a characteristic time-scale has been found.

Fig. 5 shows the structure function for the quasar sample. In order to make a clear distinction between Seyfert galaxies, which are discussed below, a limit of $z > 0.5$ and $M_B < -23$ was set for the quasars, yielding a sample of 401 objects. The errors were derived by splitting the sample into subunits and measuring the dispersion in the associated structure functions. The structure function in Fig. 5 has a near-power-law form, with a logarithmic slope of 0.20 ± 0.01. No allowance has been made for any time dilation effect since this will not affect the power-law index (Kawaguchi et al. 1998), and it is easy to verify that in a narrow redshift range the index remains the same, albeit with a somewhat larger uncertainty due to the smaller sample size.

There is no indication of a turnover in the quasar structure function at long time-scales, which confirms the impression from the light curves in Fig. 4 that the quasars have not been monitored for long enough to reach a characteristic time-scale. Equally there is no sign of the observations becoming dominated by noise at short time-scales. This again confirms an impression from Fig. 4 that the coherent variations are much larger than the errors on the observations.

4.2 Seyfert galaxies

Seyfert galaxies are well known to vary in brightness on a time-scale of months or less, which makes them relatively easy to monitor within the constraints of modern policies for the allocation of telescope time. However, the project to monitor NGC 5548 (Peterson et al. 1999) has surpassed all other efforts. Fig. 6 (top panel) shows the whole light curve sampled every 50 d for the purpose of measuring the structure function. The bottom panel shows the light curve from a particularly well-observed period in 1993–94, sampled every 10 d. Error bars are not plotted, as the errors are approximately the same size as the points of the plots.

The length and frequency of observation of this light curve make it ideal for evaluating the structure function. The observations were made with spectrographs on a number of telescopes, as spectra were essential to the investigators' main programme. The photometric observations were obtained by defining a window in the spectrum close to the V band and integrating within it. For most of the observations the errors are about 2 per cent (0.02 mag), but sometimes rise to twice this, about as large as the points in Fig. 6.

The structure functions for the two light curves in Fig. 6 are shown in Fig. 7. Errors were estimated by resampling the data in various ways and measuring the resulting dispersion in the structure functions. The left-hand panel shows a rise to longer time-scales with a gradually decreasing slope, becoming flat at a time-scale of about a year. The logarithmic slope does not become truly linear, and so to investigate the behaviour on short time-scales we refer to the right-hand panel for data sampled every 10 d. Here there is a well-defined power-law relation with logarithmic slope 0.38 ± 0.01 with an eventual flattening at longer time-scales.

The survey (Hawkins 1996) described in the previous section

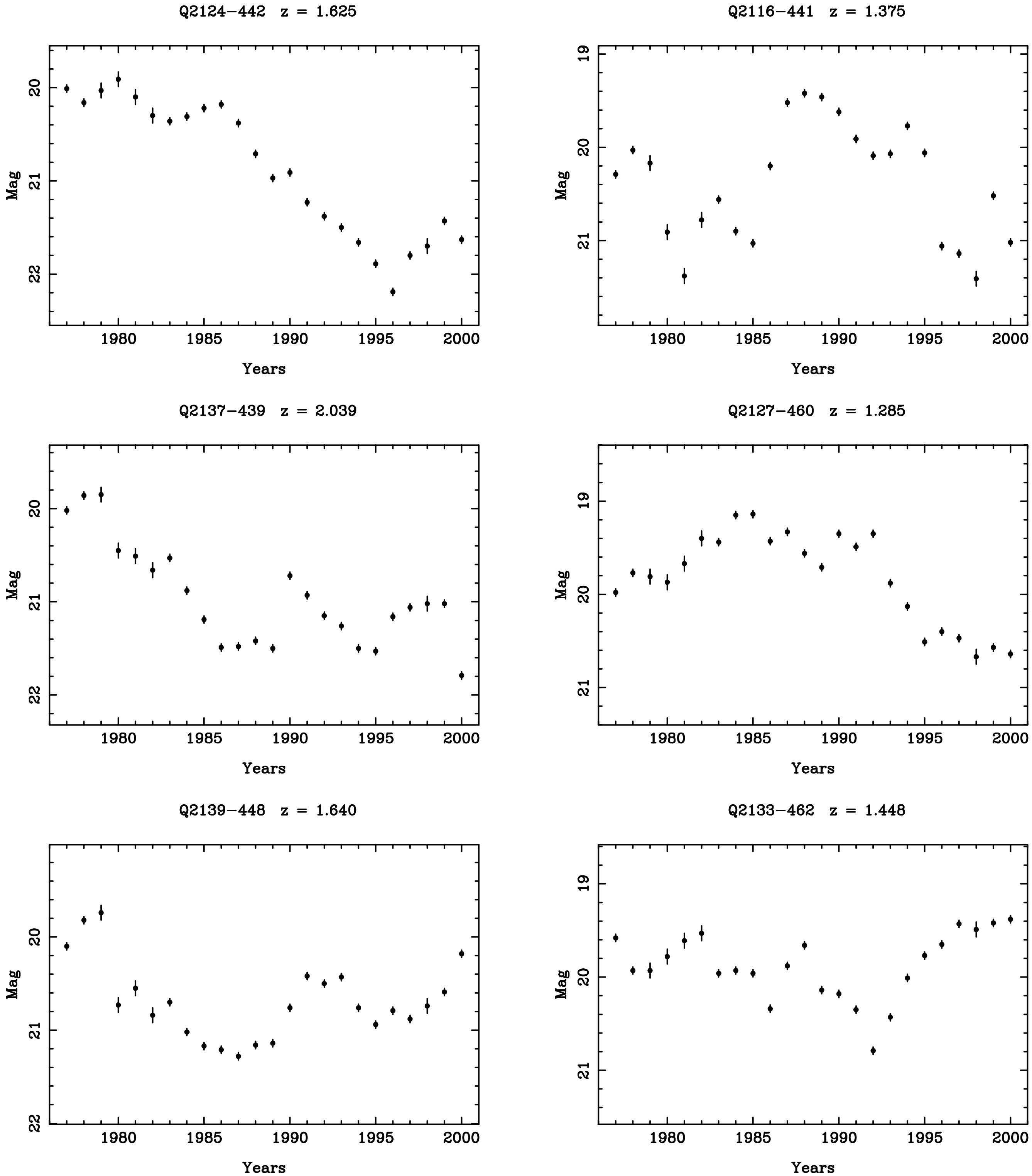

Figure 4. Light curves from the survey of Hawkins (1996), illustrating typical characteristics for quasars.

contains, in addition to the quasars, a number of Seyfert galaxies. Here again, in order to make a clear distinction with the quasar sample, we define our Seyfert galaxy sample with $M_B > -23$ and $z < 0.3$. The light curves for this sample of 45 members have a very different character from those of the quasars. Fig. 8 shows some typical examples, with errors and other details as for Fig. 4. On the whole the light curves have small amplitudes with variations characterized by small fluctuations on short time-scales, and smooth long-term gradients of apparently modest amplitude.

The structure function for this sample is shown in Fig. 9 with errors calculated as for Fig. 5. Again the data for Seyfert galaxies show a very different character from those for the quasars. The amplitude is much smaller, as would be expected from the appearance of the light curves. There is a linear section of the plot, which has a

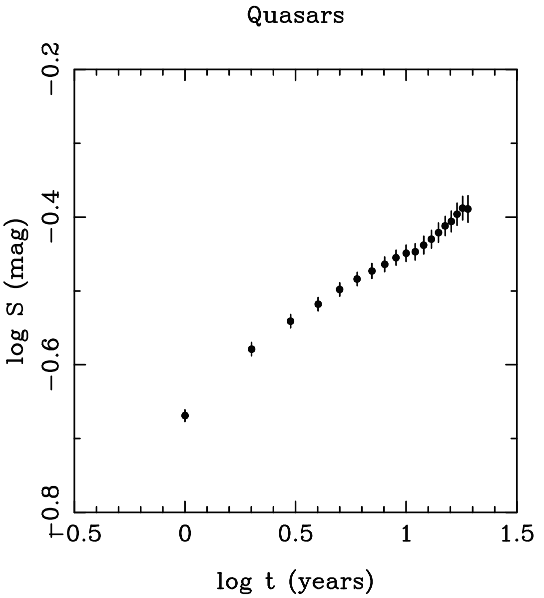

Figure 5. Structure function for the light curves of a sample of 401 quasars from the survey of Hawkins (1996).

slope of 0.36 ± 0.02, close to that of NGC 5548, and much larger than the figure of 0.20 ± 0.01 for the quasars. The structure function flattens at short time-scales. The reason for this is not immediately apparent, but may be related to measurement errors.

5 TIME ASYMMETRIES IN STRUCTURE FUNCTIONS

5.1 Model predictions

In addition to the slope of the structure function, Kawaguchi et al. (1998) discuss a further discriminant between models of AGN variability using the functions S_- and S_+ defined in Section 2 above. These functions provide a measure of the underlying asymmetry of the emission process as manifested in the light curves. It is a fortuitous circumstance that the three models for variability discussed above make qualitatively different predictions for the relationship between S_- and S_+.

The disc instability model as put forward by Kawaguchi et al. (1998) predicts a spectrum of 'avalanches' in the accretion disc which involve a gradual brightening of the disc followed by a sudden drop in flux. This pattern will manifest itself as a statistical asymmetry in the light curves, which can be detected by comparing S_- and S_+. In this case S_- should be larger than S_+ towards shorter time-scales.

By contrast, the starburst model is dominated by the light of supernovae, which are well known to rise rapidly in brightness at the onset of the explosion, after which they gradually fade. Again, this asymmetry will be apparent in the AGN light curve as a statistical asymmetry, but this time such that S_+ is larger than S_- towards shorter time-scales (Kawaguchi et al. 1998). In both cases the difference between S_- and S_+ depends on the model parameters, but should be measurable with sufficiently good data.

In the case of microlensing, which is an essentially symmetrical process in the context of quasar variability, it should not be possible to determine which way time is running by any statistical tests on the light curves. S_- and S_+ should be indistinguishable and lie on top of each other.

5.2 Observations

The idea of using time asymmetry in quasar light curves as a diagnostic for quasar variability has already received some attention in the literature (Hawkins 1996; Aretxaga 1997). The conclusion of both these studies is that statistically the light curves show no detectable departure from symmetry. The data used to

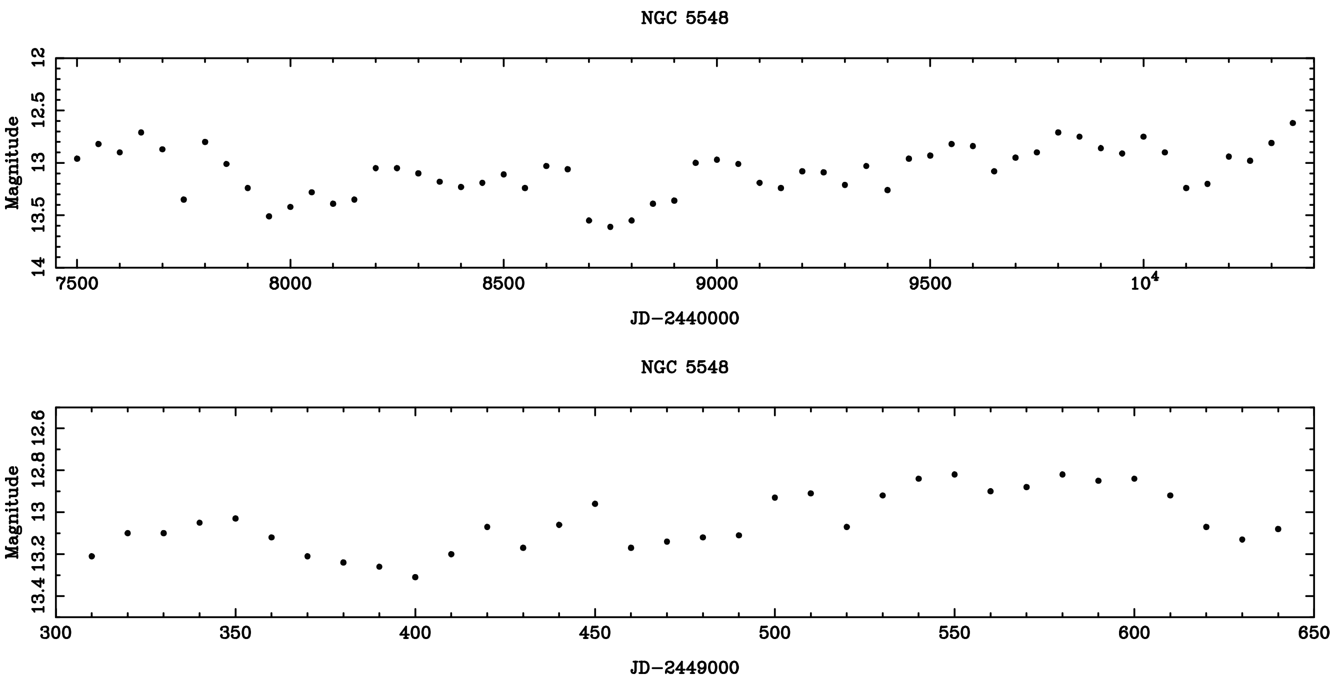

Figure 6. The light curve for NGC 5548 from Peterson et al. (1999) from 1989 December till 1996 November (top panel) and from 1993 November till 1994 October (bottom panel).

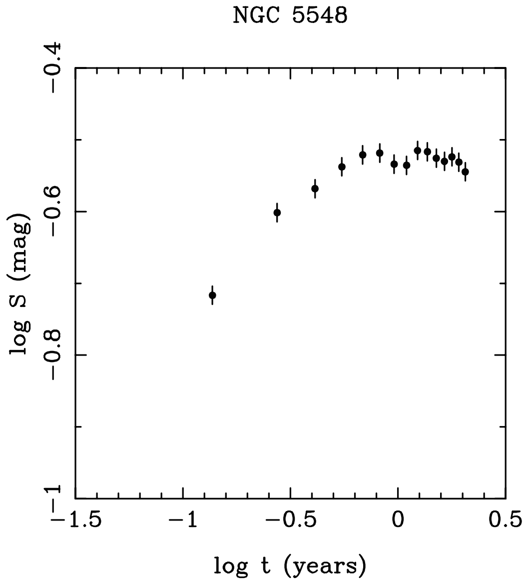

Figure 7. Structure functions for the light curve of NGC 5548 from Peterson et al. (1999) and references therein for 50-d intervals (left panel) and 10-d intervals (right panel).

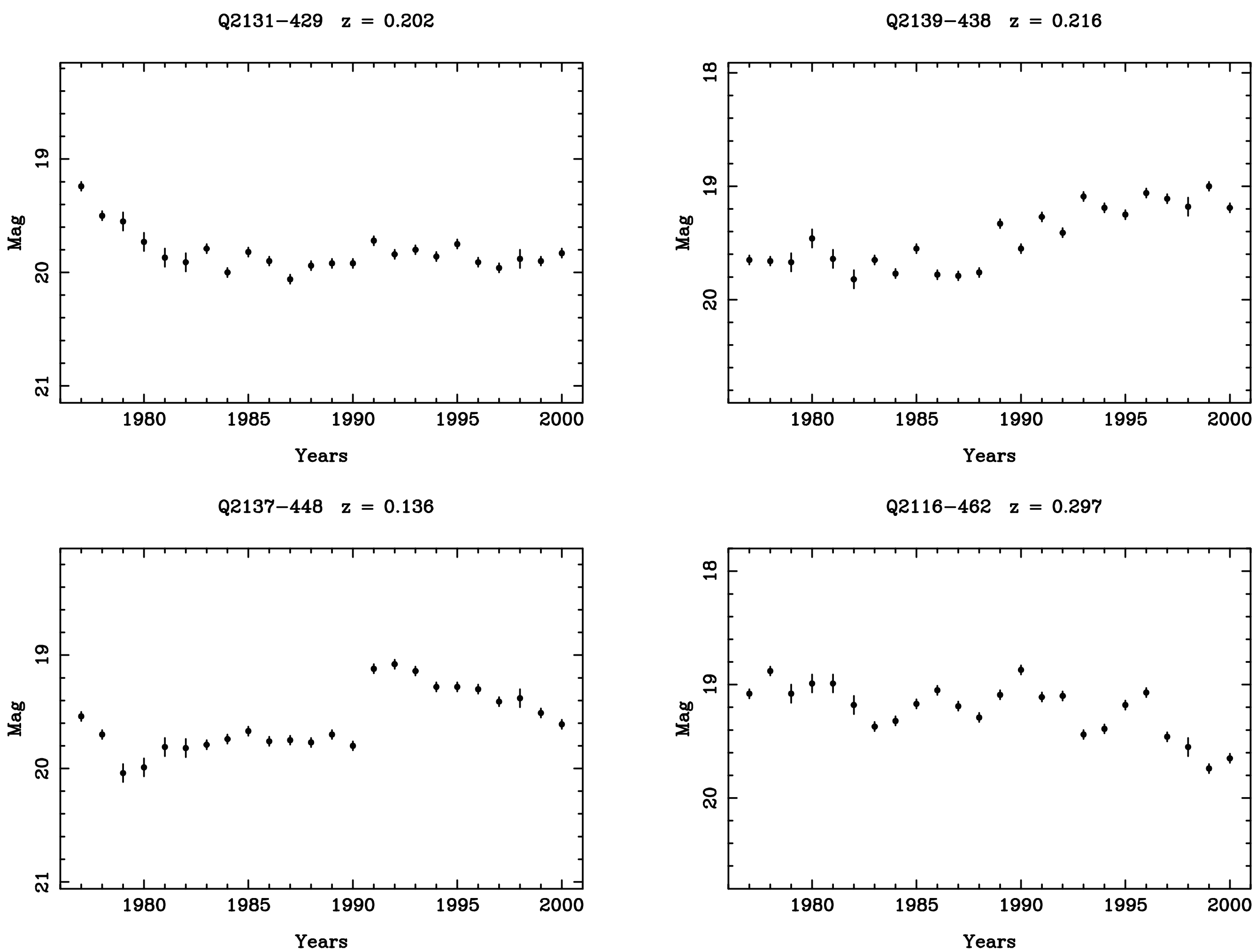

Figure 8. Light curves from the survey of Hawkins (1996), illustrating typical characteristics for low-redshift Seyfert galaxies.

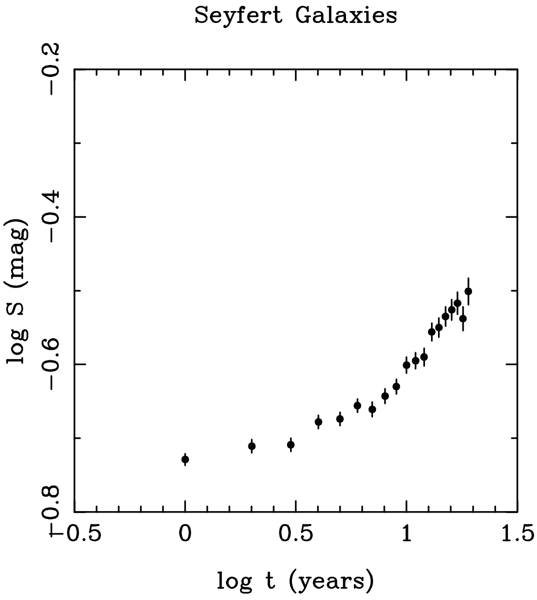

Figure 9. Structure function for the light curves of a sample of 45 Seyfert galaxies from the survey of Hawkins (1996).

produce the structure functions in Figs 5 and 7 can also be used to calculate the functions S_- and S_+. Fig. 10 shows these two functions plotted together for comparison. Error bars have been omitted in the interests of clarity, but are as for the data in Figs 5 and 7.

The S_- and S_+ functions for quasars in the left-hand panel lie very close to each other, and are indistinguishable within the errors. This implies that the observed variation is symmetrical or time-reversible. The data for the Seyfert galaxy NGC 5548 on the other hand show strong and significant differences between the two functions. On short time-scales S_- and S_+ both have a logarithmic slope of about 0.4 but S_- is systematically larger than S_+. This implies that there is an asymmetry in the light curve in the sense that the flux increases more slowly than it falls (Kawaguchi et al. 1998). On a time-scale longer than three months the two curves separate, which is probably due to the limited length of the data run with 10-d sampling.

6 CONFRONTATION OF MODELS WITH OBSERVATIONS

We are now ready to test the three models described in Section 3, which we shall do separately for quasars and Seyfert galaxies. The model predictions for the slope of the structure function are 0.83 ± 0.08, 0.44 ± 0.03 and 0.25 ± 0.03 for the starburst, accretion disc and microlensing models respectively. These figures are given in Table 1, along with the observed slopes for comparison. The slope measured for the quasar sample was 0.20 ± 0.01, which clearly rules out the starburst model. It also appears to be a long way from the accretion disc model but is consistent with the predictions for microlensing. It is however worth noting that for resolved sources the microlensing structure function gets steeper, presumably as high-frequency detail is blurred out, and there will certainly come a point where it is inconsistent with the data.

We can further compare the models by testing for time asymmetries in the light curves. For short time-scales the starburst, accretion disc and microlensing models predict $S_- < S_+, S_- > S_+$ and $S_- = S_+$ respectively. In fact, as we have seen, for quasars S_- and S_+ are effectively coincident, again favouring the microlensing model.

With Seyfert galaxies we find a very different picture from quasars. The slope of the structure function for NGC 5548 and for the Seyfert galaxy sample can be taken as 0.38 ± 0.01, which again appears to rule out the starburst model, although this time it lies close to the prediction of the accretion disc model. The microlensing model also appears to be excluded for any plausible value of the source size. Again we can further test the models by looking for time asymmetries, and we have found that for Seyfert galaxies $S_- > S_+$. This is in agreement with the prediction for the accretion disc model, and confirms the result for the structure-function slope.

To summarize, for Seyfert galaxies the observations favour the

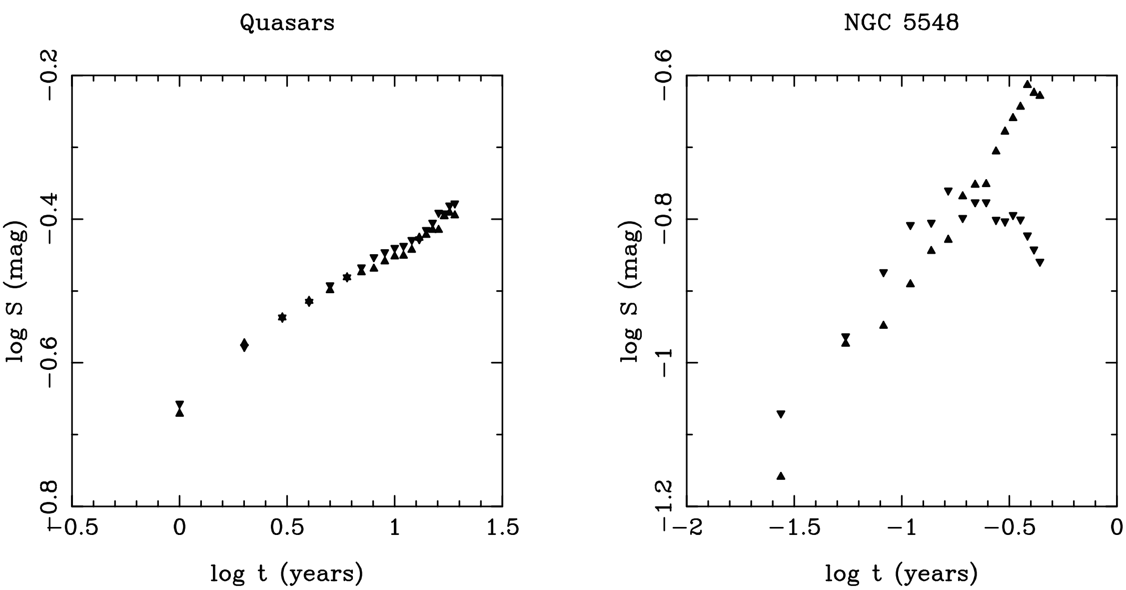

Figure 10. Time-asymmetrical structure functions for the data in Fig. 5 (left panel) and Fig. 7 (right panel). Functions for increasing (S_+) and decreasing (S_-) brightness are shown by upward- and downward-pointing triangles respectively.

Table 1. Structure-function slopes.

Model predictions	
Model	Slope
Starburst	0.83 ± 0.08
Disc instability	0.44 ± 0.03
Microlensing	0.25 ± 0.03
Observations	
AGN class	Slope
Seyfert galaxies	0.36 ± 0.02
Quasars	0.20 ± 0.01

accretion disc model, and for quasars the microlensing model. It appears that the starburst model is ruled out for all AGN.

7 DISCUSSION

Although, within the terms of reference of this paper, the results in the previous section appear rather conclusive, there are a number of caveats that must be made clear. The predictions for all three models were based on a very limited coverage of the relevant parameter spaces, and although in each case the dispersion in the structure-function slope did not appear to be large, and very much less than the differences between the mean predictions of the models, it is entirely possible that a more thorough exploration of the parameter spaces could alter the picture significantly. No doubt in due course such data will become available, and the results given here can be tightened accordingly.

In this paper we have concentrated on the slope of the structure function because for all three variability mechanisms considered it appears to be a robust quantity with little dependence on the choice of model parameters. There are other quantities such as the time-scale of variability or the amplitude of the turnover in the structure function that could in principle also be used. The problem here is the strong dependence on model parameters, or sensitivity to the window function of the data. We therefore content ourselves for the moment with the structure-function slope. However, having once decided which variability mechanism is in operation, the process can be reversed to estimate the values of model parameters.

The cellular-automaton model for disc instability (Mineshige et al. 1994) has been used here as a prototype for accretion disc models because of the testable predictions that are now available (Kawaguchi et al. 1998). There is of course much controversy over the details of emission from an AGN accretion disc, and it may well be that other accretion disc models make very different predictions for the logarithmic slope of the structure function. When testable model data become available it seems possible that the conclusions of this paper will have to be modified. It also goes without saying that the predictions of most models can be stretched to accommodate data by a careful choice of parameters, and it may be that the conclusions of this paper turn out to be less conclusive than they appear to be at present.

Notwithstanding these caveats, the conclusions reached in Section 6 appear to be well supported by the data. Taking the results at face value, we find that photometric variation in Seyfert galaxies is best explained by instabilities in an accretion disc, while for quasars the dominant mechanism for variability is microlensing. At first sight this might seem paradoxical, but in fact the two explanations sit quite well together. We can account for them by a model in which the fluctuations in the light from the accretion disc become smaller with increasing luminosity, but the effects of microlensing become more pronounced at higher redshift, which for quasars typically means higher luminosity.

This model can be tested directly in the rather rare situations where quasar variations can unambiguously be separated into those which are intrinsic to the quasar and those which are not. The best example for this purpose is the gravitational lens system 0957+561, which is a luminous quasar split into two components by the lensing effect of an intervening galaxy. The two images have been monitored intensely over the last 20 yr or so, primarily for the purpose of determining the time delay between the two components and hence measuring the value of Hubble's constant (Kundić et al. 1997). The time delay was measured from the displacement of features that occurred in both light curves and are clearly intrinsic to the quasar. These features are on the whole of short duration ($\sim$100 d) and small amplitude ($\sim$0.1 mag), as may be seen from fig. 4 of Kundić et al. (1997).

It has been known for a long time that, as well as these intrinsic variations, the two components of 0957+561 also vary independently on a time-scale of several years with an amplitude greater than 0.3 mag (Pelt et al. 1998). This mode of variation is generally accepted as microlensing by compact bodies along the line of sight to the quasar images. The difference between the two modes of variation is well illustrated in fig. 1 of Kundić et al. (1997) where the small repeating features contrast with the larger long-term variation due to microlensing. These observations support the idea that for quasars intrinsic variations are small (certainly smaller than for Seyfert galaxies) and that microlensing can easily dominate the observed variations.

8 CONCLUSIONS

In this paper we have tested the predictions of three models for variability in AGN against observations. We have compared the predicted logarithmic slope of the structure function for starburst, accretion disc instability and microlensing models with the observed structure function for Seyfert galaxies and quasars. We find that for Seyfert galaxies the starburst and microlensing models are not compatible with the observed data, but the accretion disc model is in good agreement with the data. For quasars the starburst and accretion disc instability models are ruled out, while the microlensing model agrees well with the observations.

We conclude that the variations in AGN are best explained by a model in which at low luminosities the observed variations are caused by instabilities in the accretion disc. With increasing luminosity this type of variation becomes smaller in amplitude and the observed variation becomes dominated by the effects of microlensing. This increase is associated with the larger optical depth to microlensing for most quasars.

ACKNOWLEDGMENTS

I wish to thank members of the Anglo-Australian Observatory and the Wide Field Astronomy Unit of the University of Edinburgh for their support in obtaining and measuring the Schmidt plates needed for this project.

REFERENCES

Antonucci R., 1993, ARA&A, 31, 473

Aretxaga I., 1997, Rev. Mex. Astron. Astrofis., 6, 207
Aretxaga I., Terlevich R., 1994, MNRAS, 269, 462
Aretxaga I., Cid Fernandes R., Terlevich R., 1997, MNRAS, 286, 271
Cid Fernandes R., Terlevich R., Aretxaga I., 1997, MNRAS, 289, 318
Cristiani S., Trentini S., La Franca F., Aretxaga I., Andreani P., Vio R., Gemmo A., 1996, A&A, 306, 395
George I. M., Fabian A. C., 1991, MNRAS, 249, 352
Gondhalekar P. M., Kellet B. J., Pounds K. A., Matthews L., Quenby J. J., 1994, MNRAS, 268, 973
Hawkins M. R. S., 1993, Nat, 366, 242
Hawkins M. R. S., 1996, MNRAS, 278, 787
Hawkins M. R. S., 2000, A&AS, 143, 465
Hook I. M., McMahon R. G., Boyle B. J., Irwin M. J., 1994, MNRAS, 268, 305
Kawaguchi T., Mineshige S., Umemura M., Turner E. L., 1998, ApJ, 504, 671
Kayser R., Refsdal S., Stabell R., 1986, A&A, 166, 36
Kundić T. et al., 1997, ApJ, 482, 75
Lewis G. F., Miralda-Escudé J., Richardson D. C., Wambsganss J., 1993, MNRAS, 281, 647
Mineshige S., Ouchi B., Nishimori H., 1994, PASJ, 46, 97
Pelt J., Schild R., Refsdal S., Stabell R., 1998, A&A, 336, 829
Peterson B. M. et al., 1999, ApJ, 510, 659
Pica A. J., Smith A. G., 1983, ApJ, 272, 11
Rees M. J., 1984, ARA&A, 22, 471
Schneider P., Weiss A., 1987, A&A, 171, 49
Terlevich R., Tenorio-Tagle G., Franco J., Melnick J., 1992, MNRAS, 255, 713
Trevese D., Kron R. G., Majewski S. R., Bershady M. A., 1994, ApJ, 433, 494
Wallinder F. H., Kato S., Abramowicz M. A., 1992, A&AR, 4, 79
Wyithe J. S. B., Turner E. L., 2001, MNRAS, 320, 21

This paper has been typeset from a TEX/LATEX file prepared by the author.