TESTING MICROVARIABILITY IN QUASAR DIFFERENTIAL LIGHT CURVES USING SEVERAL FIELD STARS

JOSÉ A. DE DIEGO1,2,3, JANA POLEDNIKOVA2,4, ANGEL BONGOIANI2,4, ANA M. PÉREZ GARCÍA2,4, MARIO A. DE LEO1,5, TOMÁS VERDUGO4, AND JORDI CEPÁ1,4

1 Instituto de Astronomía, Universidad Nacional Autónoma de México, Avenida Universidad 3000, Ciudad Universitaria, C.P. 04510, Distrito Federal, Mexico; jdo@astro.unam.mx
2 Instituto de Astrofísica de Canarias-Universidad de La Laguna, Spain
3 CEI Canarias: Campus Atlántico Tricontinental, E-38205 La Laguna, Tenerife, Spain
4 Departamento de Astrofísica, Universidad de la Laguna, Spain
5 Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA

Received 2014 December 10; accepted 2015 May 7; published 2015 July 15

ABSTRACT

Microvariability consists of small timescale variations of low amplitude in the photometric light curves of quasars and represents an important tool to investigate their inner core. Detection of quasar microvariations is challenging because of their non-periodicity, as well as the need for high monitoring frequency and a high signal-to-noise ratio. Statistical tests developed for the analysis of quasar differential light curves usually show either low power or low reliability, or both. In this paper we compare two statistical procedures to perform tests on several stars with enhanced power and high reliability. We perform light curve simulations of variable quasars and non-variable stars and analyze them with statistical procedures developed from the $F$-test and the analysis of variance. The results show a large improvement in the power of both statistical probes and a larger reliability when several stars are included in the analysis. The results from the simulations agree with those obtained from observations of real quasars. The high power and high reliability of the tests discussed in this paper improve the results that can be obtained from short and long timescale variability studies. These techniques are not limited to quasar variability; on the contrary, they can be easily implemented for other sources, such as variable stars. Their applications to future research and to the analysis of large-field photometric monitoring archives could reveal new variable sources.

Key words: galaxies: photometry – methods: statistical – quasars: general – techniques: photometric

1. INTRODUCTION

Variability is an important tool to study the inner physics and structure of active galactic nuclei (AGNs). The discovery of variable behavior in quasars (Matthews & Sandage 1963) closely followed the discovery of quasars themselves. The timescales of the variations in optical wavelengths range from minutes to years. In general, the amplitude of the variation is larger when observed on longer timescales and shorter wavelengths. Large amplitude and timescale variations are relatively easy to detect, but the detection becomes problematic for small amplitudes and timescales of minutes. Therefore, most, if not all, microvariability studies are performed with telescopes with apertures larger than 1 m (e.g., Kidger & de Diego 1990; Carini & Miller 1992; Gopal-Krishna et al. 1995; Ramírez et al. 2004).

We define optical microvariability as flux changes on timescales ranging from minutes to hours. The expected change in brightness for such events is on the order of hundredths of a magnitude (e.g., de Diego et al. 1998), posing a challenge for detection because the amplitude of the variation event and the noise level are similar. Moreover, quasar variability is aperiodic, and therefore difficult to analyze because we cannot collect observations to obtain a smoothly folded light curve. These limitations place a strong emphasis on careful handling of the data and the use of robust and powerful statistical techniques for the analysis. Analyzing aperiodic variations has been the subject of recent work on AGNs (Emmanoulopoulos et al. 2010; de Diego 2010, 2014, the latter is hereafter referred to as Paper I), as well as young stars and massive stars (Findeisen et al. 2015).

This issue was already recognized by several groups that proposed various statistical approaches to search for microvariability events. Thus, there are several tests that have been frequently used to report microvariations. The C-test, as proposed by Jang & Miller (1997) and Romero et al. (1999), is a very simple methodology that gained popularity 10 years ago, but unfortunately it is not a valid statistical test (de Diego 2010, Paper I). The $F$-test is also a simple statistical procedure and has been replacing the C-test in recent years. However, the original version of the $F$-test requires the quasar to be compared with a star of the same brightness, which is a condition that is difficult to meet. Howell et al. (1988) and Joshi et al. (2011) proposed scaling the photometric errors in order to work out this problem. Moreover, the $F$-test faces another limitation: lower power should be expected because of the violation of normality in data obtained from variable sources (Paper I). The one-way analysis of variance (ANOVA) was proposed by de Diego et al. (1998) for microvariability detection, and it has been shown to be superior to the $F$-test for this purpose (de Diego 2010, Paper I).

Another strategy that has been considered is the use of multiple tests (multitesting) to face the problem of the low reliability of some studies, which can be affected by the odd behavior of a comparison star rather than the variability of the target quasar. Multitesting has been implemented using two different tests or two different comparison stars (Joshi et al. 2011; Hu et al. 2013). Microvariability detection is then claimed only if both tests, at a significance level of $\alpha = 0.01$, agree on the rejection of the null non-variability hypothesis. However, in Paper I it was shown that this is a low power methodology that yields unreliable results. Nevertheless, there
is the possibility of including several comparison stars in the light curve analysis in such a way that the power of the test would be increased, as well as the reliability. In this sense, the enhanced $F$-test proposed in Paper I involves several comparison field stars in a single $F$-test, which provides a reliable methodology to increase the statistical power.

In this paper we present the nested ANOVA, which is an improved version of ANOVA that also includes several stars in the quasar differential light curve analysis. We compare enhanced $F$-test and nested ANOVA results using both simulations and examples from available real data.

The paper is organized as follows. Section 2 describes the light curve simulations, the enhanced $F$-test, and the nested ANOVA methodology. In Section 3 we show and compare the results of the analysis of the simulated light curves and the results obtained from real observations of the quasars US 3150 and 1E 15498+203. Section 4 presents a brief discussion and conclusions.

## 2. METHODS

This study consists of 1000 light curve simulations for a variable quasar of magnitude $V = 17$, a non-variable control object of the same brightness, and 21 bright field stars of magnitude $16 \leq V \leq 18$, along with a comparison of the simulations with real observations. Figure 1 shows an example of the simulations of differential light curves: a variable quasar and a non-variable control star with the same photometric error. Each differential light curve consisted of 35 observations obtained by subtracting the raw magnitudes of a reference star from the target to eliminate the changing-conditions effects between exposures. This reference star is the brightest non-saturated star in the field.

In contrast, the ANOVA procedures use only reference stars, but the test is performed between different groups of differential photometry observations of the same target rather than between the target and a comparison star’s light curves. As in Paper I, in the case of variable quasars, the light curves were the result of a random walk function with steps normally distributed with a mean of 0 and a standard deviation of 0.006 ($\mathcal{N}(0,0.006)$ RW), and a normally distributed photometric error of $\varepsilon \approx 0.007$. The photometric errors for the set of 21 field stars are also normally distributed. All the photometric errors correspond to 60 s exposures in the $V$ band with the Harold L. Johnson 1.5 m telescope at the Observatorio Astronómico Nacional in Sierra San Pedro Mártir (México), and they were calculated using the simulator for this instrument.6

As detectors become larger and detector arrays more popular, there is a tendency toward larger field sizes and thus images that include more bright stars. This situation makes it necessary to perform realistic simulations. To deal with this necessity, the magnitudes for the 21 field stars were obtained using the star distribution in the $V$ band for a galactic latitude $b^\parallel = 60^\circ$ (Allen & Cox 2000, Table 19.11, pp. 482–3). The star distribution in magnitudes was approximated using a second order polynomial:

$$\log N_V = -4.026 + 0.5607V - 0.00984V^2,$$

where $N_V$ is the cumulative function of the number of stars brighter than magnitude $V$ in one square degree of the sky. The $V$ magnitudes for the 21 reference and comparison stars were obtained applying the inversion method of the cumulative function to the uniform distribution on the $N_V$ values in the range $16 \leq V \leq 18$, i.e., up to one magnitude of difference with respect to the quasar. This range of magnitudes ensures that the images of the stars are suitable for differential photometry with the quasar, i.e., images are not suffering either overexposure or underexposure. The parameters for the stars, in the order in which they appear in the simulations, are shown in Table 1.

### 2.1. Enhanced $F$-test

The simulations were analyzed using the enhanced $F$-test, described in Paper I. Until now, the $F$-test had been implemented with a single comparison star, whose differential light curve is compared with the differential light curve obtained from a target object using the ratio of their respective variances. It is important that the mean brightness of both the comparison star and the target object are matched to ensure that the photometric errors are equal, or at least that these errors are corrected to account for the differences in brightness, as proposed by Howell et al. (1988) and Joshi et al. (2011). The enhanced $F$-test makes use of several comparisons stars, and it consists of transforming the comparison star’s differential light curves to have the same photometric noise, as if their magnitudes exactly matched the mean magnitude of the quasar under study. Note that the enhanced $F$-test needs to use a reference star in order to obtain the differential light curves.

---

6 Developed by Alan Watson, V. 1.2.0: http://www.astrossp.unam.mx/~resast/watsonBCh/fotometria-ccd.html.
Thus, the number of available bright stars to perform the test is reduced by one, and the test is performed using only the remaining bright comparison stars.

Figure 2 helps us to understand the algorithm to perform the enhanced F-test both in our simulations and in a real case. Figure 2(a) shows the magnitudes of the 21 simulated stars and their associated errors obtained with the San Pedro Martir instrument simulator. Panel (b) shows the differential light curve means and their associated errors for 20 comparison stars using the brightest star as a reference; the dashed line indicates the exponential fit to the errors. Panel (c) shows the differential light curves for the 20 stars combined in a single set of $20 \times 35 = 700$ indexed observations (the star bins of 35 observations each are identified in the upper axis); after subtracting each star mean and transforming the errors to the same level as the quasar (in our case using the exponential fitting function), the 20 differential transformed light curves are stacked (lower red light curve). Panel (d) zooms on the stacked light curve shown in the previous panel.

Thus, the number of available bright stars to perform the test is reduced by one, and the test is performed using only the remaining bright comparison stars.

Figure 2 helps us to understand the algorithm to perform the enhanced F-test both in our simulations and in a real case. Figure 2(a) shows the magnitudes of the 21 simulated stars and their associated errors obtained with the San Pedro Martir instrument simulator. Panel (b) shows the differential light curve means and their associated errors for 20 comparison stars using the brightest star as a reference; the dashed line indicates the exponential fit to the errors. Panel (c) shows the differential light curves for the 20 stars combined in a single set of $20 \times 35 = 700$ indexed observations (the star bins of 35 observations each are identified in the upper axis); after subtracting each star mean and transforming the errors to the same level as the quasar (in our case using the exponential fitting function), the 20 differential transformed light curves are stacked (lower red light curve). Panel (d) zooms on the stacked light curve shown in the previous panel.

2.2. Nested ANOVA

We also analyzed the simulations using the nested ANOVA probe. The tests of the ANOVA family require groups of replicated observations. In our case, these groups are arranged by five observations of a given quasar in a time lapse of around 5–10 minutes for telescopes with apertures larger than 1 m, so
that the quasar brightness can be considered constant for the observations collected in the same group. This is the same procedure used in previous real observations (de Diego et al. 1998; Ramírez et al. 2004, 2009) and simulations (de Diego 2010, Paper I). Of course, microvariations shorter than the lapse time within the group observations would not be detected by ANOVA, but such phenomena, known as spikes, have been seldom reported in the literature (Sagar et al. 1996; de Diego et al. 1998; Krishna et al. 2000; Stalin et al. 2004).

The nested ANOVA, as implemented in this paper, consists of the study of the variances at three different experimental levels. The first level is the one in which we are really interested: the differences between the groups, which are a signature of the quasar variability. The second level corresponds to the differences between the observations due principally to shot noise and sky subtraction. The third level accounts for the variance between the tests caused by the different reference stars. A complete description of the nested ANOVA procedure is given in Appendix A. Calculations have been performed using the R language, and the nested ANOVA probes were carried out using the aov function included in the stats package. The complete code used for the nested ANOVA analysis applied to real data can be found in Appendix B.

Note that the ANOVA family tests compare the dispersion of the individual differential magnitudes of the quasar within the groups and the dispersion between the groups, without using any comparison stars. Instead, all the bright stars are utilized as reference stars to build distinct differential light curves, and thus we will always be able to include one more star in the nested ANOVA analysis than in the enhanced F-test.

3. RESULTS

3.1. Simulations

The results in Paper I showed that ANOVA has more power than the F-test to detect variability when only one comparison star is taken into account in the analysis. Moreover, the power of the F-test was reduced by the probably non-normal distribution of the variable quasar data. This power loss can be compensated using the enhanced F-test, which includes more comparison stars in the analysis. Here we present the results of increasing the number of stars using nested ANOVA, which lead to a test power improvement, and we compare the outcomes obtained using both nested ANOVA and the enhanced F-tests.

The stars in the simulations have different magnitudes and errors, as explained in Section 2. Figure 3 shows the number of false detections or Type I errors, and the number of detections. The magnitudes and photometric errors for the stars have been obtained from a realistic star field at galactic latitude $b = 60^\circ$, as explained in the text.
1 in Figure 3 is empty for the enhanced $F$-test, but not for the nested ANOVA. The number of bright stars is directly proportional to the size of the field. For a galactic latitude $b = 60^\circ$, we expect around 0.13 stars per square arcminute, i.e., around 2 stars in a $4' \times 4'$ detector.

As the number of stars included in the analysis increases, the tests are more and more powerful, tending to an upper asymptotic value. We see in Figure 3 that in our simulations the number of detections for the enhanced $F$-test is larger than the number for the nested ANOVA for eight or more stars. The enhanced $F$-test starts with very few detections for a few and dim stars (that dominate the star distribution), but the number of detections increases dramatically and overcomes the detection for the nested ANOVA as the number of stars increases and bright stars are introduced into the analysis.

The behavior of both the nested ANOVA and the enhanced $F$-test can be explained as the result of the test designs and the data attributes. In the case of nested ANOVA, all the stars are used as reference stars, and all of them have the same weight in this probe. However, the results of the enhanced $F$-test depend very much on the photometric errors of the reference star. These errors affect the precision of the whole differential photometry. As the dim stars dominate the star distribution, we need large fields to include bright stars. We can see this effect on the enhanced $F$-test detection curves in Figures 3(b) and (d). When a bright star such as star number 6 becomes the reference star, the number of enhanced $F$-test and nested ANOVA detections is comparable, and when the even brighter stars 8 and 9 successively became the reference stars (and star 6 becomes a comparison star), the enhanced $F$-test overcomes nested ANOVA.

The inclusion of comparison stars brighter and dimmer than the quasar allows a better fit for the photometric errors as a function of the differential magnitudes. If all the stars are dimmer than the quasar, the exponential fit that we have used must be extrapolated to the quasar brightness, which is less accurate than interpolating the error. In fact, we have found that this inaccuracy produces too many Type I errors. This may be an inconvenience in the cases of few field stars or of bright quasars. If comparison stars brighter than the quasar are not available, at least for simulated data it is safer to give more weight to the star with the brightness that is more similar to that of the quasar, although this procedure reduces the power of the enhanced $F$-test. In a real situation, the errors may be individually estimated using another procedure (e.g., Howell et al. 1988; Joshi et al. 2011), but generalizing such a procedure for simulated data is beyond the scope of this paper.

As the tests with more stars are not fully independent of the test with fewer stars (for example, 14 of the 15 stars included in the 15 star tests are the same as in the tests for 14 stars), the Type I errors are not random, but show a memory of the errors obtained in the previous tests. Finally, the effect size estimate (for example, the $F$ value of the ratio of variances in the enhanced $F$-test) is fairly constant when adding more stars in the analysis. Although $F$ approaches an asymptotic value and the probabilities decline smoothly with the number of stars (Figure 4(b)), for a given light curve the analysis may yield fluctuations in the probabilities obtained with different numbers of stars. In those cases when the $F$ value is close to the critical limit of detection, the tests may fluctuate between detections and non-detections as shown in Figure 4(c).

![Figure 4](image_url)

**Figure 4.** Two examples of the enhanced $F$-test output. Panel (a) shows the $F$ value calculated as the variance ratio between the simulated quasar light curve and the stacked star light curves for a clearly variable quasar (filled squares) and a less variable one (hollow circles). Panel (b) shows the logarithm of the probabilities that the data from the earlier quasar have been drawn from a non-variable object. The probabilities show a decreasing trend, as expected from the power increase when more stars are added in the analysis. Panel (c) is like panel (b) but for the later quasar. Although the $F$ values show very small variations for this object and the probabilities show the expected trend as in panel (b), the tests are just in the limit of detection at the 0.001 level (dashed line) and tend to fluctuate between detections and non-detections.

### 3.2. Observations

Data discussed in this section correspond to the radio-quiet quasars US 3150 and 1E 15498+203, and have been already reported in de Diego et al. (1998). Here we present a new analysis including more field stars to compare the results obtained with nested ANOVA and the enhanced $F$-test. The quasar light curves are shown in Figure 5. Observations of US 3150 ($V \approx 16.8$) were obtained during about 6 hr of monitoring on 1996 November 13, with the f/13.05 1.5 m telescope at San Pedro Martir Observatory (Baja California, Mexico) equipped with a Thomson THX 31156 CCD of 1024 $\times$ 1024 pixels of 19 $\times$ 19 $\mu$m and Metachrome II coated. Quasar US 3150 showed clear microvariability at a level of significance of $\alpha = 0.001$ and an amplitude of $\Delta m \approx 0.1$ mag.
The data consist of seven groups of five observations with exposures of 60 s in the $V$ band. There were five stars in the 3.4 arcmin field suitable for our purpose, but at the time when we performed the observations, we had no plans to use all these stars in our analysis. Hence, the brightest star (more than 3 mag brighter than the quasar) was overexposed in the last two groups of observations. Therefore, only for demonstration purposes, we have dropped these two groups and we performed the analysis using the first five sets of observations, corresponding to about 4 hr of monitoring.

Observations of 1E 15498+203 ($V \approx 17.12$) were obtained during 4.5 hr of monitoring on 1997 May 12 with the same telescope, but this time equipped with a Tektronix TK1024AB of 1024 $\times$ 1024 pixels of 24 $\times$ 24 $\mu$m, thinned and Methylchrome II coated. Quasar 1E 15498+203 showed only weak evidence of variability at $\alpha = 0.01$, an amplitude of $\Delta m \approx 0.06$ mag. The data consist of five groups of five observations with exposures of 60 s in the $V$ band. There were four stars in the 4.3 arcmin field suitable for our purpose, but photometry was affected by the scattered light from the bright star HD 142109 ($V = 8.94$) at around 1$'$ from the quasar, which spoils the observations of some stars in the third group. Consequently, we have dropped this group from our analysis.

Tables 2 and 3 show the results obtained with the enhanced $F$-test and nested ANOVA for US 3150 and 1E 15498+203, respectively. In both tables, column 1 indicates the number of stars used in the analysis. Columns 2 and 3 indicate the $F$ value and the associated probability for the enhanced $F$-test. Columns 4 and 5 indicate the same values for the nested ANOVA probe. Columns 6 and 7, and 8 and 9, indicate the same values for the control star. The number of stars in the enhanced $F$ tests is one less than in the corresponding nested ANOVA probes because the brightest star was used as a reference. The number of stars in the control tests is one less than in the target quasar because the star most similar in brightness to the quasar was used as a control.

Figure 5. Monitored light curves for quasars US 3150 (1996 November 13) and 1E 15498+203 (1997 May 12) and a comparison star. Panels (a) and (b) show the US 3150 raw and binned differential light curves (upper points) and a comparison star (lower points). Error bars in panels (b) and (d) correspond to the standard error measured within each group of observations.
is the light
\[ \alpha = \text{all the tests yield results below the significance level} \]
both the quasars and the control stars. In the case of US 3150, the nested ANOVA, the number of degrees of freedom is always 3 for the numerator and 16 for the denominator. The number of degrees of freedom for the enhanced

| KK | KK |
|----|----|
| 4  | 2.477 4.6E-3 5.808 7.0E-3 |
| 3  | 2.384 1.2E-2 7.198 2.8E-3 0.963 0.52 0.467 0.71 |
| 2  | 3.098 1.2E-2 5.231 1.0E-2 1.252 0.32 0.361 0.78 |
| 1  | ...

Note. The number of degrees of freedom for the enhanced \( F \)-test is 19 for the numerator and 17, 36, and 55 for the denominator, for 1, 2, and 3 stars, respectively. For the nested ANOVA, the number of degrees of freedom is always 3 for the numerator and 16 for the denominator.

3.2.1. Enhanced \( F \)-test

We used the brightest star as a reference to build the photometric differential light curves. To perform the enhanced \( F \)-test it is necessary that the errors of the comparison stars are transformed to the same level as the errors of the quasar. In principle, any function that reasonably fits the differential light curve standard deviation and the mean magnitude relationship can be used. However, in another work (J. Polednikova et al. 2015, in preparation) we have fitted exponential curves to a large set of comparison stars in larger CCD fields and find that such a function accurately describes this relationship. The function is of the form:

\[
s_i = s_0 + A e^{\mu},
\]

where \( s_i \) is the light curve standard deviation and \( \mu_i \) is the light curve mean magnitude for the \( i \)th star, while \( s_0 \) and \( A \) are the parameters to be fitted. Note that fitting two parameters to the comparison star data reduces in two units the number of degrees of freedom for the comparison stars of the enhanced \( F \)-test. Figure 6 shows the relationship between the differential light curve standard deviation for the comparison stars and the quasars (excluded from the fit), and their mean differential magnitudes. The exponential curve fits the data for the comparison stars up to a precision of \( \approx 0.001 \) mag.

The data for the quasars shown in Figure 6 are clearly far away from stars of comparable brightness and the fitted exponential curve, indicating that their standard deviation may be dominated by variability rather than photometric errors. Yet, the \( F \)-test with a single comparison star fails to detect variations in US 3150 at the level of significance \( \alpha = 0.01 \), and four comparison stars are needed to reach \( \alpha = 0.001 \) (Table 2). In the case of 1E 15498+203, the enhanced \( F \)-test only reaches the level of significance \( \alpha = 0.01 \) when using three stars. The \( F \) statistics for the enhanced \( F \)-test on both quasars shows small fluctuations but a monotonic decrease in the \( p \) value. Finally, the enhanced \( F \)-test for the control stars also shows small fluctuations (in US 3150) for both the \( F \) statistics and the \( p \) values, but in no case do the tests show any evidence of variation.

3.2.2. Nested ANOVA

Results of the nested ANOVA probes presented in Tables 2 and 3 show fluctuations for the \( F \) statistics and the \( p \) values for both the quasars and the control stars. In the case of US 3150, all the tests yield results below the significance level \( \alpha = 0.001 \), but for 1E 15498+203 the \( p \) values are only below \( \alpha = 0.01 \). In both cases, the control stars do not show any evidence of variations.

The results obtained by both the enhanced \( F \)-test and nested ANOVA agree that US 3150 shows strong evidence of microvariability, while for 1E 15498+203 the evidence is not conclusive, as reported in de Diego et al. (1998). They also agree with the results from the simulations, in the sense that nested ANOVA has more power to detect microvariations than the enhanced \( F \)-test when there is a limited number of available field stars.

Figure 6. Light curve standard deviations vs. mean differential magnitude dependence. The standard deviation of the light curves yields error estimates. Panel (a) shows four comparison stars (squares) in the quasar US 3150 (circle) field that have been used to fit an exponential curve to the relationship between the standard deviation and the mean differential magnitude (dashed line). Panel (b) is like the previous panel but for three comparison stars (squares) in the field of the quasar 1E 15498+203 (circle). These figures show that the data dispersion for the quasar was much larger than the dispersion for their respective comparison stars, indicating possible variations.
4. CONCLUSIONS

In this paper we have presented the nested ANOVA test and have compared the power of this probe with the enhanced F-test presented in Paper I. These tests use several field stars to analyze the light curve of a target quasar. Both the enhanced F-test and nested ANOVA are very powerful in detecting photometric variations in differential light curves, and they asymptotically tend to larger power values than their respective single star counterparts (the F-test and ANOVA). The nested ANOVA probe shows superior performance when the number of available bright stars is limited, but the enhanced F-test surpasses nested ANOVA when there are several bright stars, some of them brighter than the target object. This is a consequence of using a bright reference star with the enhanced F-test that diminishes the errors of the differential photometry.

In our simulations, the power increases with the number of stars included in the analyses from approximately 40% to 65% for the nested ANOVA probes, and from 5% to 90% for the enhanced F-test, both at a significance level of $\alpha = 0.001$. These results agree with those obtained from the observed light curve of quasars US 3150 and 1E 15498+203, for which the probabilities tend to lower values as the number of stars increases, while nested ANOVA yields lower probabilities than the enhanced F-test due to the limited number of field stars. In this paper, we have demonstrated that the power of microvariability studies increases significantly when including several stars in the analysis, and that this procedure is an important improvement with respect to all the other tests that have been used in the quasar microvariability literature, from the C-test to the single star F-test and ANOVA, as well as multi-testing techniques.

From these results, we conclude that both the enhanced F-test and the nested ANOVA probe are robust techniques to detect quasar microvariability. They not only add power to the microvariability detection, but by introducing several stars in the light curve analysis, ensure more reliable results than those obtained when using a single star. In small fields with a limited number of stars, we encourage the use of nested ANOVA rather than the enhanced F-test because under these circumstances nested ANOVA is more powerful, and it can include one more bright star in the analysis.

The use of both the enhanced F-test and nested ANOVA in future research will improve the detection of microvariability in quasars. This improvement will eventually yield a better understanding of the physics involved in microvariability phenomena and the central quasar’s engine. The enhanced F-test can be easily applicable in selecting varying source candidates (i.e., quasars and variable stars) in regions visited several times during large photometric surveys, such as the the Legacy Survey Stripe 82 from the Sloan Digital Sky Survey. The wide field images in such surveys, which include a large number of stars, allow us to build a mean reference star with sources different from the stars used for comparison in the enhanced F-test, thus ensuring the independence of the reference mean star and the comparison stars.

This research has been supported by the UNAM-DGAPA-PAPIIT IN110013 Program and the Canary Islands CIE: Tricontinental Atlantic Campus. This work is based on observations acquired at the Observatorio Astronómico Nacional in the Sierra San Pedro Mártir (OAN-SPM), Baja California, Mexico. J.P., A.B., A.M.P.G., and J.C. are grateful for the support from the grant AYA2011-29517-C03-01 of the Spanish Ministerio de Economía y Competitividad (MINECO). M.A.D. L. thanks the UC MEXUS-CONACYT Postdoctoral Fellowship in the UCR Department of Physics and Astronomy. The authors are thankful to Jana Benda for the linguistic revision and to the anonymous referee for the constructive suggestions.

APPENDIX A
NESTED ANOVA

The nested ANOVA analysis presented in this paper involves simulations of $a = 7$ groups, each of them consisting of $b = 5$ images (35 in total) of a target quasar and a number $n \leq 21$ of field stars that can be used to replicate measurements of the quasar differential photometry. In practice, the images contained in a given group were obtained in a time lapse that is short (less than 10 minutes) in comparison with the microvariability timescale of the quasar, while the gap between groups of observations that will be compared is arbitrary.

This methodology implies $n$ pseudoreplicates per image to yield a total of $35n$ readings of the quasar differential photometry. We use the term pseudoreplicates for the $n$ readings of a given image to note that these readings are not independent because we use the same quasar observation (the same image) to estimate the quasar differential photometry with $n$ stars. Images are the experimental units of our study: as we have 35 images (34 degrees of freedom in total), no matter how many stars we use in our analysis, the degrees of freedom are divided into 6 for the (7) groups of images and 28 ($=34 - 6$) for errors. The number of replicate measurements that we perform on a given image will improve the precision on that image, and the variability among the images from a given group yields a measurement of the variability within that group.

Figure 7 shows the schematic representation of the observational methodology: an arbitrary number of groups of observations (seven in our example), five images of the quasar in each group, and $n$ differential photometry measurements from each image, using $n$ different reference stars. In fact, this is a two-stage nested design, probably the simplest nested ANOVA. Following Montgomery (2013, chapter 14), the linear statistical model is expressed by:

$$y_{ijk} = \mu + \gamma_i + \omega_{j(i)} + \epsilon_{ijk},$$

where $y_{ijk}$ corresponds to the measurement of the quasar differential photometry using the reference star $k = 1, 2, \ldots, n$, located in the image $j = 1, 2, \ldots, b$ of the $i = 1, 2, \ldots, a$ group of observations. The true mean of the quasar light curve is denoted by $\mu$ and the deviation from the true mean of the $i$th group is denoted by $\gamma_i$. The deviation of the $j$th image of the quasar with respect to the mean of the $i$th group is denoted by $\omega_{j(i)}$, where the subscript $j(i)$ indicates that the $j$ observation is nested under the $i$th group. Finally, $\epsilon_{ijk}$ corresponds to the error term.

The model described by Equation (3) yields a sum of (deviation) squares that, after some manipulation
(Montgomery 2013), can be expressed as
\[
\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y})^2 = bn \sum_{i=1}^{a} (\bar{y}_i - \bar{y})^2 \\
+ n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij} - \bar{y}_i)^2 \\
+ \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ijk})^2,
\]
where the horizontal line over a letter indicates the mean value (e.g., \(\bar{y}_i = (\sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk})/bn\)). Equation (4) can also be expressed symbolically as
\[
SS_T = SS_G + SS_{G(i)} + SS_E,
\]
where \(SS_T\) denotes the total sum of squares, \(SS_G\) the sum of squares due to groups, \(SS_{G(i)}\) the sum of squares due to the nested observations in groups, and \(SS_E\) the residual sum of squares due to errors. The degrees of freedom for each square sum is \(abn - 1\) for \(SS_T\), \(a - 1\) for \(SS_G\), \(a(b - 1)\) for \(SS_{G(i)}\), and \(ab(n - 1)\) for \(SS_E\). Dividing a sum of squares \(SS\) by the respective degrees of freedom \(\nu\) yields the mean square \((MS = SS/\nu)\). The ratios between mean squares are distributed as \(F\).

In our simulations, the groups of observations are drawn from an infinite pool of instants in the continuous light curve of a target object, and the observations in each group are drawn from an infinite pool of possible images, and thus both groups and observations constitute random effects. In this case we assume that \(\gamma_i\) is \(\text{NID}(0,\sigma^2)\) and \(\omega_{j(i)}\) is \(\text{NID}(0,\sigma^2)\), where \(\text{NID}(0,\sigma^2)\) is normally and independently distributed with mean 0 and variance \(\sigma^2\). Then, we test \(H_0: a^2 = 0\) with \(MS_G/MS_{G(i)}\) and \(H_0: \sigma^2 = 0\) with \(MS_{G(i)}/MS_E\).

The residuals for a two-stage nested design are given by
\[
e_{ijk} = y_{ijk} - \bar{y}_{ijk},
\]
where \(\bar{y}_{ijk}\) are the individual image averages, and the fitted value is
\[
y_{ijk} = y + \hat{\gamma}_i + \hat{\omega}_{j(i)}.
\]

For random effects, the analysis of variance method allows us to estimate the variance components \(\sigma^2, \sigma^2_G, \) and \(\sigma^2\) using the expected mean squares:
\[
\begin{align*}
\bar{\sigma}^2 &= MS_E \\
\bar{\sigma}^2_G &= MS_{O(G)} - MS_E \\
\bar{\sigma}^2 &= MS_G - MS_{O(G)}/bn.
\end{align*}
\]

### APPENDIX B

**NESTED ANALYSIS OF VARIANCE R CODE**

```r
# Nested ANOVA
# (c) J. A. de Diego—November 2014

# ************************
# Input file example (file without headers)

# TIME QUASAR S1 S2 S3 S4 S5 GROUP
03:08:50 21.504 17.389 17.851 18.911 21.187 21.577 1
03:10:44 21.557 17.347 17.882 18.951 21.232 21.655 1
03:12:24 21.488 17.400 17.840 18.913 21.207 21.597 1
03:14:12 21.481 17.319 17.844 18.908 21.178 21.593 1
03:15:54 21.484 17.348 17.841 18.919 21.176 21.577 1
03:21:24 21.601 17.354 17.877 18.938 21.232 21.611 2
03:33:29 21.552 17.338 17.872 18.930 21.222 21.597 2
04:34:54 21.635 17.368 17.895 18.966 21.217 21.647 2
04:36:19 21.583 17.334 17.897 18.952 21.233 21.607 2
04:37:44 21.581 17.363 17.894 18.963 21.240 21.651 2

# The number of columns for the stars is arbitrary
# Magnitudes for the quasar and stars are as generated by IRAF
# routines, i.e., unprocessed (not differential or extinction corrected)
# Columns TIME and GROUP are not used in this code, but probably needed
# for figures and other computations.
# This code assumes that the last star corresponds to the control.
```

**NESTED ANOVA**

- Stars A, B, C, etc., are images that include the quasar and a number of stars. We can obtain a differential photometry measurement of the quasar from each star.
- The residuals for a two-stage nested design are given by
  \[ e_{ijk} = y_{ijk} - \bar{y}_{ijk}, \]
- where \(\bar{y}_{ijk}\) are the individual image averages, and the fitted value is
  \[ y_{ijk} = y + \hat{\gamma}_i + \hat{\omega}_{j(i)}. \]
Allocate matrices
# ****************************

doqlc <- colnames(qlc)
qlc <- g

for(j in 2:(s))
  # Perform test
  prob.anova.qn[j-1] <- summary(aov(ym ~ tm))
  prob.anova.qv[j-1] <- nest.anova.sum.qv[(j-1)]

# Test and probabilities for the quasar
nest.anova.sum.qv[(j-1)] <- summary(aov(ym ~ tm))
prob.anova.qv[j-1] <- nest.anova.sum.qv[(j-1)][1][1][1]["Pr(>F)][[1]"

# Temporal control variables
ym <- as.vector(tapply(dlcj$qn, list(dlcj$s, dlcj$ao), mean))

# Test and probabilities for the quasar
nest.anova.sum.qn[(j-1)] <- summary(aov(ym ~ tm))
prob.anova.qn[j-1] <- nest.anova.sum.qn[(j-1)][1][1]["Pr(>F)][[1]"

# Remove last element for the control star results (it is a test with itself)
nest.anova.sum.qn[(j-1)] <- NULL
prob.anova.qn <- prob.anova.qn[~(j-1)]

#---------------------
# Allocate matrices
#---------------------

nest.anova.sum.qn <- vector("list")
nest.anova.sum.qv <- vector("list")

prob.anova.qv <- matrix(, nrow = 1, ncol = (s-1))
colnames(prob.anova.qv) <- paste('cs.', 1:(s-1), sep="")

prob.anova.qn <- matrix(, nrow = 1, ncol = (s-1))
colnames(prob.anova.qn) <- paste('cs.', 1:(s-1), sep="")

# Perform test
for(j in 2:(s))
  # Temporal quasar variables
dlcj <- doqlc[doqlc$st < j]

ym <- as.vector(tapply(dlcj$qv, list(dlcj$s, dlcj$ao), mean))

#---------------------
# Read data from "mydata.dat" and prepare variables
#---------------------

test <- read.table("US3150_V3.dat")
s <- length(test) - 2 # Number of sources in the field (including the quasar)

n <- dim(test)[1] # Data points in the light curve
ng <- 5 # Number of points in each group

g <- n/ng # Number of groups

doqlc <- test
colnames(doqlc) <- c("Time", "q", paste0("s", seq(1:(s-1))), "Group")
dqlcj <- list(dlcj)
dqlcj$tm <- as.vector(tapply(as.numeric(dlcj$s),
list(dlcj$s, dlcj$ao), mean)) # vector of observation (image) means
dqlcj$tm <- factor(doqlc$tm)

doqlc$tm <- factor(doqlc$tm)

doqlc$tm <- factor(doqlc$tm)

doqlc$tm <- factor(doqlc$tm)

#---------------------
# References
#---------------------

REFERENCES

Allen, C., & Cox, A. 2000, in Allen’s Astrophysical Quantities, ed. A. N. Cox
(New York: Springer)

Carini, M. T., & Miller, H. R. 1992, ApJ, 385, 146

de Diego, J. A. 2010, AJ, 139, 1269

de Diego, J. A. 2014, AJ, 148, 97

de Diego, J. A., Dultzin-Hacyan, D., Ramirez, A., & Benitez, E. 1998, ApJ, 501, 69

Emmanoulopoulos, D., McHardy, I. M., & Uttley, P. 2010, MNRAS, 404, 931

Finkeisen, K., Cody, A. M., & Hillenbrand, L. 2015, ApJ, 798, 89

Gopal-Krishna, Sagar, R., & Wiita, P. J. 1995, MNRAS, 274, 701

Howell, S. B., Warnock, A., III, & Mitchell, K. J. 1988, AJ, 95, 247

Hu, S. M., Chen, X., & Guo, D. F. 2014, JApA, 35, 261

Jang, M., & Miller, H. R. 1997, AJ, 114, 565

Joshi, R., Chand, H., Gupta, A. C., & Wiita, P. J. 2011, MNRAS, 412, 2717

Kidger, M. R., & de Diego, J. A. 1990, A&A, 251, L25

Krishna, G., Gupta, A. C., Sagar, R., et al. 2000, MNRAS, 314, 815

Kurtz, D. C. 2013, Design and Analysis of Experiments, Vol. 17
(ed.; Hoboken, NJ: John Wiley & Sons Inc.), 730

Ramírez, A., de Diego, J. A., Dultzin, D., & González-Pérez, J. N. 2009, AJ, 138, 30

Stalin, C. S., Gopal-Krishna, Sagar, R., & Wiita, P. J. 2004, MNRAS, 350, 175

de Diego, J. A., Dultzin-Hacyan, D., Ramirez, A., & Benitez, E. 1998, ApJ, 501, 69

Joshi, R., Chand, H., Gupta, A. C., & Wiita, P. J. 2011, MNRAS, 412, 2717

Kidger, M. R., & de Diego, J. A. 1990, A&A, 227, L25

Krishna, G., Gupta, A. C., Sagar, R., et al. 2000, MNRAS, 314, 815

Matthis, T. A., & Sandage, A. R. 1963, ApJ, 138, 30

Montgomery, D. C. 2013, Design and Analysis of Experiments, Vol. 17 (8th ed.; Hoboken, NJ: John Wiley & Sons Inc.), 730

Ramírez, A., de Diego, J. A., Dultzin, D., & González-Pérez, J. N. 2009, AJ, 138, 30

Ramírez, A., de Diego, J. A., Dultzin-Hacyan, D., & González-Pérez, J. N. 2004, A&A, 421, 83

Romero, G. E., Cellone, S. A., & Combi, J. A. 1999, A&AS, 135, 477

Sagar, R., Gopal-Krishna, & Wiita, P. J. 1996, MNRAS, 281, 1267

Stalin, C. S., Gopal-Krishna, Sagar, R., & Wiita, P. J. 2004, MNRAS, 350, 175