Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **453,** 1701–1726 (2015) doi:10.1093/mnras/stv1718

Simulations of the OzDES AGN reverberation mapping project

Anthea L. King,[1,2⋆] Paul Martini,[3,4] Tamara M. Davis,[1] K. D. Denney,[3,4] C. S. Kochanek,[3,4] Bradley M. Peterson,[3,4] Andreas Skielboe,[2] Marianne Vestergaard,[2,5] Eric Huff,[3,4] Darach Watson,[2] Manda Banerji,[6] Richard McMahon,[6] Rob Sharp[7] and C. Lidman[8]

[1]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[2]*Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark*
[3]*Department of Astronomy, Ohio State University, 140 W. 18th Ave,Columbus, OH 43221, USA*
[4]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[5]*Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[6]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[7]*Research School of Astronomy & Astrophysics, Australian National University, Canberra, ACT 2611, Australia*
[8]*Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia*

Accepted 2015 July 25. Received 2015 July 25; in original form 2015 April 12

ABSTRACT

As part of the Australian spectroscopic dark energy survey (OzDES) we are carrying out a large-scale reverberation mapping study of $\geq$500 quasars over five years in the 30 deg^2 area of the Dark Energy Survey (DES) supernova fields. These quasars have redshifts ranging up to 4 and have apparent AB magnitudes between 16.8 mag $< r <$ 22.5 mag. The aim of the survey is to measure time lags between fluctuations in the quasar continuum and broad emission-line fluxes of individual objects in order to measure black hole masses for a broad range of active galactic nuclei (AGN) and constrain the radius–luminosity (R–L) relationship. Here we investigate the expected efficiency of the OzDES reverberation mapping campaign and its possible extensions. We expect to recover lags for ∼35–45 per cent of the quasars. AGN with shorter lags and greater variability are more likely to yield a lag measurement, and objects with lags $\lesssim$6 months or ∼1 yr are expected to be recovered the most accurately. The baseline OzDES reverberation mapping campaign is predicted to produce an unbiased measurement of the R–L relationship parameters for H β, Mg II λ2798, and C IV λ1549. Extending the baseline survey by either increasing the spectroscopic cadence, extending the survey season, or improving the emission-line flux measurement accuracy will significantly improve the R–L parameter constraints for all broad emission lines.

Key words: quasars: general – quasars: supermassive black holes – dark energy.

1 INTRODUCTION

There is good evidence that supermassive black holes (SMBHs) are present at the centre of all massive galaxies (e.g. Kormendy & Richstone 1995; Richstone 1998; Ferrarese & Ford 2005), and that there are tight, empirical relationships between the mass of the SMBH and properties of the host galaxy, such as stellar velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000; McConnell & Ma 2013), light concentration (Graham et al. 2001), and bulge luminosity and stellar mass (Richstone 1998; Kormendy & Gebhardt 2001; Marconi & Hunt 2003; McConnell & Ma 2013). These relationships suggest an interplay between black hole growth and galaxy evolution; however, the true nature of this relationship is still unknown and is a major area of research in understanding galaxy evolution (e.g. King 2003, 2005; Di Matteo, Springel & Hernquist 2005; Ferrarese & Ford 2005; Murray, Quataert & Thompson 2005; Di Matteo et al. 2008; Park et al. 2015). To better comprehend the origin and evolution of the SMBH–galaxy relationship and the growth of SMBHs over cosmic time, it is necessary to obtain accurate and precise measurements of black hole masses. Direct measurements of black hole masses through stellar or gas dynamics require high spatial resolution and are therefore limited to the local Universe (Ferrarese & Ford 2005).

Active galactic nuclei (AGN) provide an alternative method of black hole mass measurement. Continuum emission from the

⋆ E-mail: anthea.king@uqconnect.edu.au

accretion disc is absorbed by gas deep within the gravitational potential of the black hole. The broad-line region (BLR) gas reprocesses this radiation and emits Doppler broadened emission lines. The emission-line luminosity varies in response to changes in the continuum emission in a roughly linear fashion with an associated time lag, τ, which is the mean light travel time from the accretion disc to the BLR, at the responsivity-weighted mean distance, $R = c\tau$. The measurement of this time lag, through detailed comparison of the emission line and continuum flux variations, is referred to as reverberation mapping (RM; Blandford & McKee 1982; Peterson 1993).

If the BLR gas is in virial equilibrium and its motion is dominated by the gravity of the SMBH, the mass of the black hole is

$$M_{\rm BH} = \frac{f c \tau \Delta V^2}{G}, \quad (1)$$

where G is the gravitational constant, τ is the measured reverberation time lag, ΔV is the line-of-sight velocity dispersion of the BLR gas estimated from the emission-line width of the RMS variance spectrum, and f is a dimensionless virial factor that converts the measured line-of-sight virial product into the true black hole mass. The virial factor depends on the geometry, kinematics, and orientation of the BLR, and although it is of the order unity, it is expected to differ between quasars.

RM has yielded lags for approximately 60 AGN (Bentz & Katz 2015). The lags exhibit a tight power-law relationship with the continuum luminosity, λL_λ (Kaspi et al. 2000; Bentz et al. 2009a, 2013), as predicted from simple photoionization physics (Davidson 1972; Krolik & McKee 1978). This strong correlation is the basis for single-epoch black hole mass estimates using a single epoch of spectroscopy (e.g. Laor 1998; Wandel, Peterson & Malkan 1999; McLure & Jarvis 2002; Vestergaard & Peterson 2006), and enables black hole masses to be estimated for a far larger sample of AGN than is possible with the full RM method (e.g. Vestergaard 2004; Vestergaard & Osmer 2009; Shen et al. 2011; Kelly & Shen 2013).

The single-epoch mass estimation technique has been used widely in cosmology (e.g. Vestergaard et al. 2008; Vestergaard & Osmer 2009; Schulze & Wisotzki 2010; Willott et al. 2010; Mortlock et al. 2011; Trump et al. 2011; Shen & Liu 2012; Kelly & Shen 2013). As a consequence, it is important to accurately and precisely determine the *R–L* relationship, and any physical properties on which its calibration depends, because both random and systematic uncertainties in the *R–L* relationship are transferred into uncertainties in the single-epoch masses and all subsequent studies. The origin of the scatter around the *R–L* relationship has been investigated by several authors (Bentz et al. 2009a, 2013; Watson et al. 2011; Kilerci Eser et al. 2015), and recent evidence suggests an additional dependence on the *R–L* relationship with the Eddington ratio (Du et al. 2015).

Additionally, it may be possible to reverse the *R–L* relationship and use the time lag to infer the intrinsic luminosity of an AGN, and therefore its luminosity distance (Watson et al. 2011). The resulting distance measurements can be used to independently probe the acceleration of the Universe and dark energy. Despite 'dark energy' appearing to be the dominant energy component of the Universe (e.g. Blake et al. 2011a,b; Conley et al. 2011; Anderson et al. 2012; Padmanabhan et al. 2012; Hinshaw et al. 2013; Planck Collaboration XVI 2014), its nature is unknown and gaining additional understanding of its properties remains a high priority. The high luminosity and prevalence of AGN would make them valuable probes of the expansion history of the universe over a greater redshift range than other methods, and thus valuable for investigating the time evolution of dark energy (Czerny et al. 2013; King et al. 2014).

Availability and allocation of telescope time and instrumentation have driven most previous RM campaigns to focus their efforts on monitoring small numbers of AGN (∼10), using relatively small (1–3- m class) telescopes for relatively small time-scales (typically <1∼3 yr; e.g. Clavel et al. 1991; Robinson 1994; Wanders et al. 1997; Collier et al. 1998; Peterson et al. 1998, 1999, 2002; Kaspi et al. 2000, 2007; Bentz et al. 2009b; Denney et al. 2010; Barth et al. 2011; Rafter et al. 2011, 2013; Grier et al. 2012; Du et al. 2014). As a consequence, they have focused on the apparently brightest and most variable objects, leading to a bias towards local, low-luminosity AGN. Both higher redshift and higher luminosity quasars have longer lags due to the effects of time dilation and the *R–L* relation, and higher luminosity quasars also have lower variability amplitudes (Vanden Berk et al. 2004; MacLeod et al. 2010), making a lag recovery less likely. While there have also been long, multiyear campaigns (up to 8 yr; Peterson 1999; Kaspi et al. 2000, 2007; Peterson et al. 2002), they only monitored a small number of quasars so the bias towards the brightest and most variable objects remains. Mitigating this bias in the current RM sample requires larger telescopes and longer multi-object observation campaigns. This paper investigates a large-scale RM campaign being run as part of the ongoing Dark Energy Survey (DES), in conjunction with the OzDES spectroscopic survey. This five-year campaign covers a large range in magnitude and redshift, allowing RM studies of an AGN sample with a broader range of properties over a large redshift range.

DES is an optical survey predominantly aimed at understanding the expansion of the Universe using four complementary methods: Type Ia supernovae (SNe), baryon acoustic oscillations (BAO), weak lensing, and galaxy cluster counts. DES officially began in the second half of 2013, and the plan is to image 5000 deg^2 with five filters (g, r, i, z, Y) over 5 yr. The supernova (SN) component of the survey will consist of repeated observations of 30 deg^2 of sky in the g, r, i, z filters, divided into two deep and eight shallow SNe fields, to detect and monitor SN and other transients.

OzDES is the leading spectroscopic counterpart to DES. It will repeatedly monitor the DES SNe fields using the Two Degree Field (2dF) multi-object fibre spectrograph (AAOmega; Saunders et al. 2004) at the Anglo-Australian Telescope (AAT). Its main science goal is to measure the redshifts of Type Ia SNe host galaxies. In addition, a number of fibres in each field will be dedicated to monitoring a select group of quasars to perform RM. OzDES, in conjunction with DES, will monitor ≥500 quasars for the full five years of the survey over the redshift range $0 < z \lesssim 4$. This is approximately a 10-fold increase in number and 2-fold increase in redshift range over the existing RM sample, and comparable in number to the ongoing Sloan Digital Sky Survey (SDSS) RM project (Shen et al. 2015).

While the OzDES RM programme will monitor hundreds of quasars, the likelihood that this programme will successfully recover reverberation lags depends on the frequency and accuracy of the light-curve measurements, the length of the survey, and the intrinsic variability of the monitored AGN sample. OzDES is expected to target each field approximately 25 times over the five-year period. This number of observation epochs is significantly smaller than traditional RM campaigns (e.g. Peterson et al. 2002; Bentz et al. 2009a; Denney et al. 2009; Barth et al. 2011), which have found that emission-line lag recovery generally requires 30–50 well-spaced epochs of observations, and favourable continuum flux variations. Our current study aims to investigate the expected efficiency of the

Table 1. DES SNe fields.

Target name	RA (h m s)	Dec. (° ′ ″)	Type
E1	00:31:29.9	−43:00:34.6	Shallow
E2	00:38:00.0	−43:59:52.8	Shallow
S1	02:51:16.8	+00:00:00.0	Shallow
S2	02:44:46.7	−00:59:18.2	Shallow
C1	03:37:05.8	−27:06:41.8	Shallow
C2	03:37:05.8	−29:05:18.2	Shallow
C3	03:30:35.6	−28:06:00.0	Deep
X1	02:17:54.2	−04:55:46.2	Shallow
X2	02:22:39.5	−06:24:43.6	Shallow
X3	02:25:48.0	−04:36:00.0	Deep

Note. RA and Dec. are given for J2000.

OzDES RM campaign by generating realistic AGN light curves and attempting to recover the input lags. We will then use our findings to determine how to optimally select our target AGN sample, and make predictions about the scientific results for our sample. We also investigate ways to improve the programme design and execution, such as increased cadence, changes to the survey length, and improved measurement accuracy, as a means to maximize the scientific output.

The outline of the paper is as follows: a technical summary of the DES and OzDES surveys is given in Section 2, followed by a description of the survey simulation in Section 3. The predictions of the efficiency of the survey are presented in Section 4, along with the expected improvements for several possible survey extensions. We examine the predicted scientific results in Section 5. Finally, the results are summarized in Section 6.

2 DES/OZDES SURVEY

2.1 Fields

The DES SNe fields were chosen to have extensive past observation histories and to overlap with the Visible and Infrared Survey Telescope for Astronomy (VISTA) survey area. The 10 chosen fields were the Elias fields (E1, E2), SDSS Stripe 82 (S1, S2), the *Chandra* Deep Fields (C1, C2, C3), and the *XMM* Large Structure Survey fields (X1, X2, X3). The coordinates of the fields are given in Table 1.

2.2 Target selection

The quasar candidates were initially chosen from (1) known quasars in the DES SNe fields with $m_{\rm r,\,psf} < 21.2$ mag [best-fitting point-spread function (PSF) magnitude], (2) point sources with $m_{\rm r,\,psf} < 21$ mag selected through the KX method[1] (Warren, Hewett & Foltz 2000) using data from DES and the VISTA Hemisphere Survey (VHS; McMahon et al. 2013; Sutherland et al. 2015), and (3) point sources with $m_{\rm r,\,psf} < 21$ mag selected through photo-z template fitting using DES, VHS, and *Wide-field Infrared Survey Explorer* (Wright et al. 2010) photometry. The details of this selection process are described in Banerji et al. (2015). We obtained spectra of 3331 quasar candidates in 2013, and after visual inspection, the sample was reduced to the current 989 objects. Fig. 1 shows the redshift and r-band magnitude distribution of this sample. We have ranked the sources based on the quality of the spectra and the number of emission lines present in the spectra. The highest priority objects are shown in blue. If it is necessary to decrease this sample to the minimum of 50 per field, we will incorporate the results of the simulations presented here into our target selection criteria and also preference the most variable quasars (for which lag recovery is most probable) based on the first two years of data.

[1] This method selects quasars based on excess flux in the K band relative to stars, which is due to the power-law nature of the quasar spectral energy distribution.

2.3 Photometry

DES uses the DECam instrument (Flaugher et al. 2010) on the Blanco 4-metre telescope at the Cerro Tololo Inter-American Observatory in Chile. Images of the 10 SNe fields are planned to be taken approximately every 5–7 d between September 1 and February 15 every year between 2013 and 2017. Photometry from 2012 November–2013 February were also taken as part of the DES Science Verification period. For a given field, images are taken in all four filters in the same night when possible (presently this occurs on 80 per cent of nights). Otherwise, images in the remaining filters are taken during the next available night. The approximate observation period of DES is shown in Fig. 2, along with the visibility of the 10 DES SNe fields. During every night of DES imaging, the SNe fields that have not been observed in the last five nights are given highest priority, with special preference given to the deep fields, C3 and X3. To date, the median gap between consecutive observations is 6.5 d and the maximum gap ranges between 12 and 21 d. For these simulations we assume that photometric observations are taken every 7 d. The nominal exposure times and corresponding limiting magnitudes for both the deep and shallow fields are given in Table 2.

2.4 Spectroscopy

The OzDES spectroscopic observations are being taken with the AAOmega spectrograph (Saunders et al. 2004; Smith et al. 2004) fed by the 2dF multi-object system (Taylor, Cannon & Watson 1997; Lewis et al. 2002) on the AAT. The 2.1 deg diameter field of view of 2dF is almost identical to that of DECam, making it the ideal instrument for the spectroscopic follow-up of DES targets (Yuan et al. 2015), and OzDES will run over a similar period of time as DES. The 2dF multi-object system is a robotic fibre positioner that allows simultaneous observations of up to 392 targets anywhere within the field of view. The projected fibre diameter of the instrument is approximately 2 arcsec. The 2dF fibres feed AAOmega, a double beam spectrograph with a wavelength coverage of 3750–8900 Å and a resolution of $R \sim 1500$.

During each run, OzDES integrations will be 2 h long in each DES field. Over the five years of the survey, a total of 100 nights have been allocated for OzDES with a graduated allocation plan. A larger number of nights have been allocated each year as the survey progresses: 12 nights in 2013, 16 in 2014, 20 in 2015, 24 in 2016, and 28 in 2017. It takes approximately four nights to observe all 10 DES SN fields (including a 33 per cent allowance for bad weather). We expect that each field will be visited approximately 25 times over the five-year period, and the cadence of the measurements will be approximately monthly within each year. The 2014 observations are now complete and each field has five or more epochs of measurements. To date, we have been able to devote fibres to 100 AGN per field to help optimize our final sample selection. This number will drop to approximately 80 per field in 2015 and may drop to a minimum of 50 per field in future years pending internal review within the OzDES collaboration.

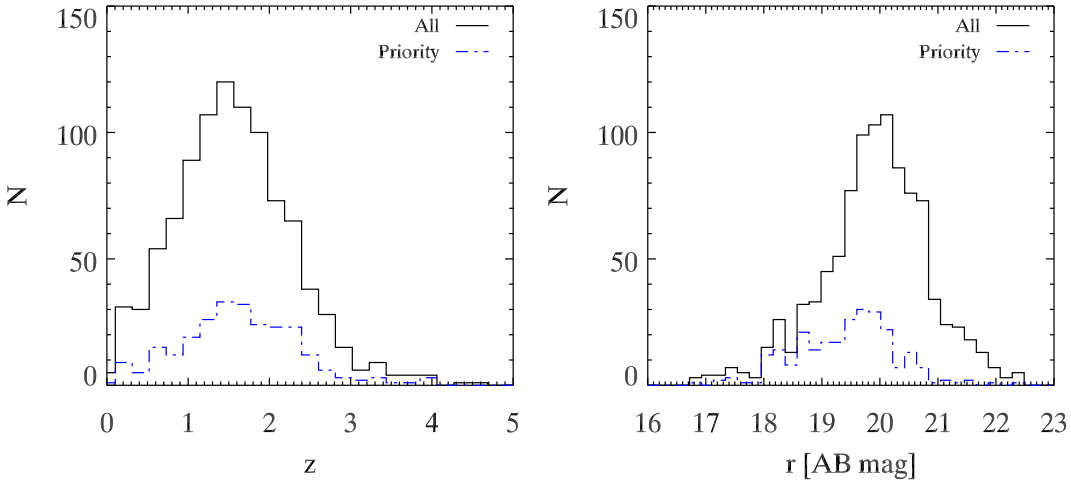


Figure 1. The redshift (left) and magnitude (right) distributions of the current OzDES quasar sample. The priority sample consists of the highest quality candidates based on visual inspection of the spectra.

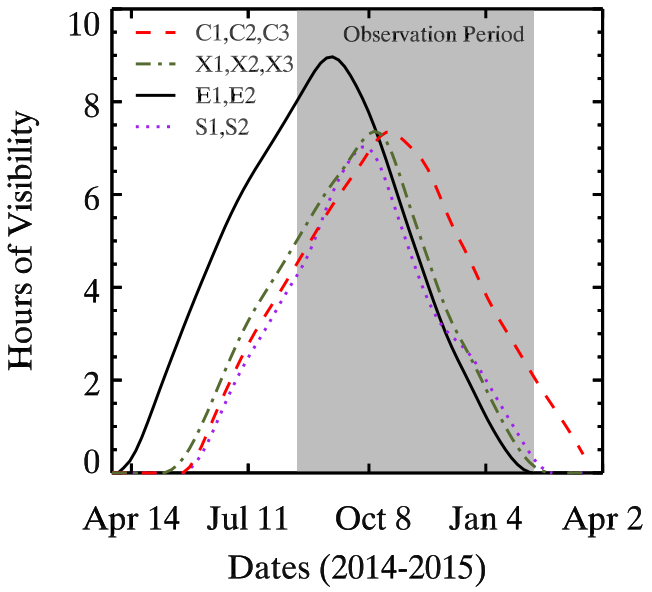


Figure 2. Number of hours the DES SN fields are visible throughout the year with an airmass of <2. The values are calculated for the period 2014 April–2015 April using JSKYCALC[2]. The shaded observation period roughly represents the time when photometric and spectroscopic data will be taken with the current DES and OzDES programme design. Section 4.2 investigates the improvements afforded by extending this observation period to fully encompass the time when the fields are visible.

The quality of the current OzDES spectra can be probed from the distribution of preliminary line flux signal-to-noise ratio (SNR) measurements shown in Fig. 3. The median SNR value for the total sample is $\mathrm{SNR_{all}} \sim 11$ and for the priority sample is $\mathrm{SNR_{priority}} \sim 21$. More details on the calculation of these preliminary SNR values are given in Appendix A. OzDES is also monitoring 10–15 F-type stars per field. The F-type star observations are used for relative spectrophotometric calibration, which is expected to be better than 10 per cent based on results from the GAMA project (Hopkins et al. 2013).

2 http://www.dartmouth.edu/~physics/labs/skycalc/flyer.html

Table 2. DES exposure times and limiting magnitudes of the SN fields.

Filter	Shallow field Exposure time (s)	Shallow field Limiting mag (AB)	Deep field Exposure time (s)	Deep field Limiting mag (AB)
g	175	24.9	600	25.6
r	150	24.3	1200	25.4
i	200	23.9	1800	25.1
z	400	23.8	3630	24.8

3 SIMULATION SETUP

Our ability to accurately recover lags is highly dependent on the presence of prominent features in the continuum and emission-line light curves. Accurate detection and characterization of light-curve features depends on (1) the frequency of measurements, (2) the length of the survey, (3) the accuracy of the measurements, and (4) the intrinsic variability of the object. In order to predict how well OzDES will be able to recover lags, we analysed mock catalogues of quasars with realistic continuum and emission-line light curves. Below we describe our methods for simulating and subsequently recovering lags from such light curves.

3.1 Mock catalogue

We began by constructing mock catalogues of 520 AGN distributed uniformly in 40 redshift bins and 13 magnitude bins over the range $0 < z < 4.0$ and $18.0 < r < 20.5$. This range roughly corresponds to the capabilities of the AAOmega spectrograph within the framework of the OzDES programme design. The small number of AGN brighter than 18th magnitude will be automatically targeted and are not considered for this investigation.

3.2 Monochromatic continuum luminosity estimation

The observed *R–L* relationship for each broad emission line has been constructed empirically using the monochromatic luminosity of a nearby continuum region as a proxy for the ionizing luminosity. The most common emission-line–continuum region pairs used for constructing *R–L* relationships, are H β with $L_{5100\,\text{Å}}$ (e.g. Peterson & Horne 2004; Vestergaard & Peterson 2006; Bentz et al. 2009a,

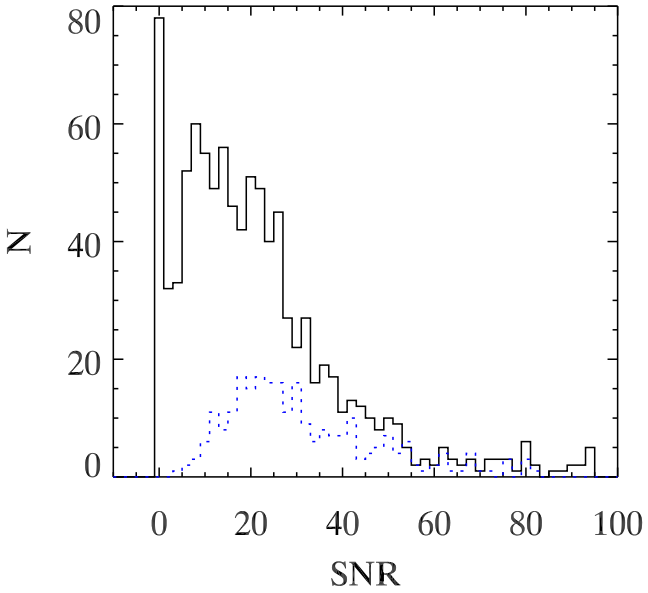


Figure 3. Distribution of emission-line signal-to-noise ratios (SNR) in the current OzDES quasar sample for the full and priority sample (defined in Section 2.2). The emission-line SNR values were estimated for H β, Mg II, and C IV for each object at all observed epochs using the OzDES spectra. The SNR assigned to each object is the median SNR value of the emission line with the highest median SNR value, based on the multiple spectral measurements available.

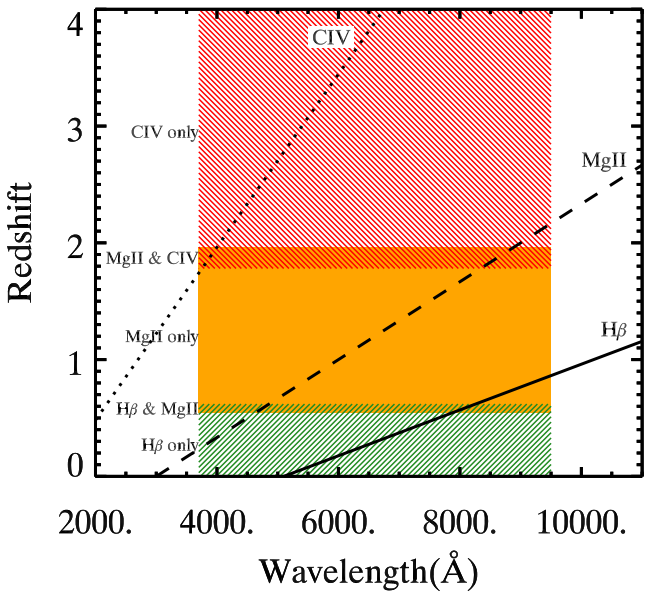


Figure 4. The five redshift regimes in which OzDES is sensitive to different broad lines (or pairs of lines) given the properties of the AAOmega spectrograph. The black curves show the observed frame wavelength of each broad emission line considered in our simulations as a function of redshift.

2013), Mg II with $L_{3000\,\text{Å}}$ (Vestergaard & Osmer 2009; Trakhtenbrot & Netzer 2012), and C IV with $L_{1350\,\text{Å}}$ (Vestergaard 2002; Peterson et al. 2004; Vestergaard & Peterson 2006; Kaspi et al. 2007; Vestergaard & Osmer 2009; Park et al. 2013; Trevese et al. 2014), and we use these pairs in our analysis. As we are restricted to the optical observing window (3750–8900 Å), the redshift range is broken into five different sections according to the observable emission lines: $z < 0.54$, H β only; $0.54 < z < 0.62$, H β and Mg II; $0.62 < z < 1.78$, Mg II only; $1.78 < z < 1.96$, Mg II and C IV; and $z > 1.9$ C IV only (see Fig. 4 for illustration). Accurate measurement of the nearby continuum level is required to measure the line flux correctly. Therefore, these redshift ranges allow a generous amount of continuum emission on either side of the broad line, and avoid bluer wavelengths where the spectrograph's throughput is low. However, our choice of redshift ranges is somewhat conservative with respect to how far into the blue and/or red each emission line or continuum region can be measured. Out of our currently observed 989 sample, the number of quasars that fall within the $18.0 < r < 20.5$ mag range and the H β only, H β and Mg II, Mg II only, Mg II and C IV, and C IV only wavelength ranges are 39, 13, 401, 76, and 231, respectively.

The monochromatic flux densities at rest frame 1350, 3000, and 5100 Å for each AGN are estimated using the known DECam *r*-band magnitudes and redshifts, and a *K*-correction based on the filter response curves of the DECam *r*-band filter and the SDSS quasar template (Vanden Berk et al. 2001). The SDSS quasar template spans the rest-frame wavelength range 800–8555 Å and is based on a sample of quasars that covers a similar redshift and magnitude range to that being observed by OzDES, so it should be a representative template for this work. The quasar template is scaled to the input magnitude of each target under the assumption that the bolometric luminosity scales as $L_{\text{bol}} = 9.0\lambda L_{\lambda}(5100\ \text{Å})$ (Kaspi et al. 2000). This bolometric correction is the approximate mid-point between the results of Richards et al. (2006) and Krawczyk et al. (2013) who found values of 10.3 ± 2.1 and 7.79 ± 1.69, respectively. The luminosity is then estimated assuming a Λ cold dark matter (ΛCDM) cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\text{M}} = 0.3$, and $\Omega_{\Lambda} = 0.7$.

This process does not take into account host galaxy contamination, Galactic extinction, or intrinsic variations in the AGN spectral energy distribution (SED). Nevertheless, it provides a sufficient luminosity estimate for our present purposes. When host galaxy contamination is included in the simulations, no significant change in the results was observed because the loss in sensitivity due to contamination from the host galaxy is generally balanced by the increased variability expected for fainter objects (see Section 3.4). For more discussion on this point see Section 6.

3.3 Lag estimation

We calculate the time lags associated with each source from the monochromatic continuum luminosity and previously published *R–L* relationships for H β, Mg II, and C IV. We adopt the H β *R–L* relationship,

$$\log_{10}(R_{\text{H}\beta}\ \text{light-days}) = -21.2 \pm 2.2 + 0.517 \pm 0.033 \times \log_{10}(\lambda L_{\lambda}(5100\ \text{Å})\ \text{erg s}^{-1}), \quad (2)$$

presented by Bentz et al. (2009a), which is derived from RM measurements of 35 AGN spanning four orders of magnitude in luminosity. Switching to the Bentz et al. (2013) relation would not appreciably change our results. The C IV *R–L* relationship is not as well determined as the H β *R–L* relationship because C IV lags have only been measured for a few objects (Koratkar & Gaskell 1991; Peterson & Horne 2004; Kaspi et al. 2007). Most of these objects have similar luminosities, and there is only a single object at each of the low- and high-luminosity ends. The C IV *R–L* relationship we use,

$$\log_{10}(R_{\text{C IV}}\ \text{light-days}) = -23.3 \pm 2.6 + (0.55 \pm 0.04) \times \log_{10}(\lambda L_{\lambda}(1350\ \text{Å})\ \text{erg s}^{-1}), \quad (3)$$

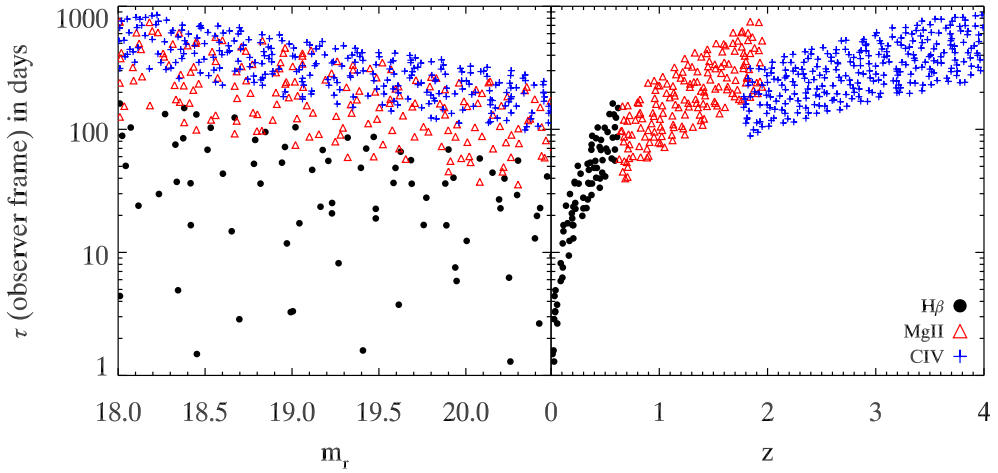


Figure 5. The distribution of lags in a mock AGN sample, as a function of *r*-band magnitude (left) and redshift, z (right). The different colours and symbols correspond to the different emission lines (black circles: H β, red triangles: Mg II, blue crosses: C IV).

is taken from Kaspi et al. (2007) and is based on only seven objects. Mg II is not yet well studied with RM, and there are only a few campaigns in which the time lag has been measured (Reichert et al. 1994; Dietrich & Kollatschny 1995; Metzroth, Onken & Peterson 2006, though the last two papers only present marginal detections). However, because the Mg II lag measurement was found to be consistent with the H β lag in two of the studies (Reichert et al. 1994; Stirpe et al. 1994; Bentz et al. 2006, using Metzroth et al. 2006), there is a strong correlation between the width of Mg II and H β emission lines[3] and since H β and Mg II have similar ionization parameters, it is generally assumed that the two lines originate at the same radius from the ionizing source. Trakhtenbrot & Netzer (2012) use this assumption to estimate a Mg II *R–L* relationship of

$$\log_{10}(R_{\rm Mg\,II}\,{\rm light-days}) = -25.72 \pm 0.62 + (0.615 \pm 0.014) \times \log_{10}\left(\lambda L_{\lambda}(3000\,{\rm \AA})\ {\rm erg\ s^{-1}}\right). \quad (4)$$

This estimate was created using an empirical correlation between $\lambda L_{\lambda}(5100\,{\rm \AA})$ and $\lambda L_{\lambda}(3000\,{\rm \AA})$, and an existing H β *R–L* relationship. As a consequence, it is less certain than an *R–L* relationship derived from direct RM measurements, but it is our only option in the absence of such measurements for Mg II.

The resulting distributions of the observed time lags (including time dilation) with redshift and magnitude are shown in Fig. 5. The C IV lags tend to be a factor of $\sim$2 smaller than the Mg II lags for the same objects. This is consistent with the findings of Kaspi et al. (2007).

3.4 Light-curve estimation

We model the AGN continuum light curves as a damped random walk (DRW) characterized by a time-scale, $\tau_{\rm D}$,[4] and the asymptotic amplitude of the structure function over long time-scales, $S_F(\infty)$. Zu et al. (2013) show this model is a good representation of quasar variability on the time-scales of the OzDES survey. Our ability to recover the time lag for each AGN is highly dependent on these variability parameters. If $S_F(\infty)$ is large and $\tau_{\rm D}$ is short, it is more likely that significant variations in the light curve will be seen during the observation campaign.

MacLeod et al. (2010) found that the values of $\tau_{\rm D}$ and $S_F(\infty)$ scale with the luminosity of the AGN, the observed wavelength, and the mass of the black hole, following the power law

$$\log_{10}(\alpha) = A_{\alpha} + B_{\alpha}\log_{10}(\lambda/4000) + C_{\alpha}(M_i + 23) + D_{\alpha}\log_{10}(M_{\rm BH}), \quad (5)$$

where α refers to either $\tau_{\rm D}$ (in days) or $S_F(\infty)$ (in mag), λ (in Å) is the continuum wavelength of interest (1350, 3000, or 5100 Å), the coefficients are

$$A_{\tau_{\rm D}} = 2.4,\ \ B_{\tau_{\rm D}} = 0.17,\ \ C_{\tau_{\rm D}} = 0.03,\ \ D_{\tau_{\rm D}} = 0.21,$$
$$A_{S_F} = -0.51,\ \ B_{S_F} = -0.48,\ \ C_{S_F} = 0.13,\ \ D_{S_F} = 0.18,$$

M_i is the absolute magnitude of the quasar and is calculated using the known *r*-band magnitude and the *K*-corrections prescribed in Richards et al. (2006) for SDSS, and $M_{\rm BH}$ is the mass of the black hole. The black hole masses were randomly assigned using the probability distribution described by MacLeod et al. (2010),

$$P(\log_{10} M_{\rm BH}|M_i) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\log_{10} M_{\rm BH} - \log_{10}\overline{M}_{\rm BH})^2}{2\sigma^2_{M_{\rm BH}}}\right], \quad (6)$$

where $\log_{10}\overline{M}_{\rm BH} = 2.0 - 0.27M_i$ and the spread in the distribution, $\sigma_{M_{\rm BH}} = 0.58 + 0.011M_i$. The mass of the black hole is in solar units. The DES and SDSS *gri* filters are quite similar, so the use of the Richards et al. (2006) *K*-corrections are justified (Honscheid et al. 2008). Additionally, we assume the same cosmological model as MacLeod et al. (2010), so the absolute magnitude scales are identical.

3.4.1 *Continuum light curve*

The continuum light curve, in magnitude, is defined as a combination of a mean μ and variable term $\Delta C(t_i)$, such that

[3] However, this correlation is not a tight one-to-one relationship between the Mg II and H β widths (Vestergaard & Wilkes 2001; Trakhtenbrot & Netzer 2012), as is commonly reported in the literature.

[4] Note that MacLeod et al. (2010) use τ without the subscript for this parameter, but we use $\tau_{\rm D}$ following Zu, Kochanek & Peterson (2011) so as not to confuse this parameter with the reverberation lag, which we denote as τ.

$C(t) = \mu + \Delta C(t)$. The value of μ is defined as the monochromatic continuum flux density at given wavelength (5100, 3000, and 1350 Å; Section 3.2) converted to a magnitude. For a DRW, the variable component is constructed by initializing the light curve at t_0 as $\Delta C(t_0) = \sigma G(1)$ where $G(1)$ is a Gaussian random deviate of unit dispersion, and $S_F(\infty) = \sqrt{2}\sigma$. Subsequent points are created using the recursion formula (see Kelly, Bechtold & Siemiginowska 2009; Kozłowski et al. 2010; MacLeod et al. 2010)

$$\Delta C(t_{i+1}) = \Delta C(t_i)\exp(-|t_{i+1} - t_i|/\tau_{\rm D}) + \sigma[1 - \exp(-2|t_{i+1} - t_i|/\tau_{\rm D})]^{1/2} G(1). \quad (7)$$

3.4.2 *Emission-line light curve*

The emission-line light curve is the response to the continuum light curve. If we only consider the temporal response of the overall emission line, the resulting light curve is given by

$$\Delta L(t) = \int \Psi(\tau)\Delta C(t - \tau){\rm d}\tau, \quad (8)$$

where $\Delta L(t)$ is the emission-line light-curve flux relative to its mean value, $\Delta C(t)$ is the variable component of the continuum light curve defined above, and $\Psi(\tau)$ is the transfer function, which describes the emission-line response to a delta function outburst in the continuum (Blandford & McKee 1982). The transfer function is related to the overall structure and kinematics of the BLR and its true form remains an active area of research (Pancoast, Brewer & Treu 2011; Zu et al. 2011; Pancoast et al. 2012; Grier et al. 2013; Skielboe et al. 2015). For simplicity we have chosen a top hat transfer function, mirroring JAVELIN (Zu et al. 2011), given by

$$\Psi(t) = \begin{cases} \frac{A}{2w} & \text{for } \tau - {\rm w} < {\rm t} < \tau + {\rm w} \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where A is a scaling term and w is the half-width of the top hat function. We set the scaling term, $A = 1$, and the half-width to $w = 0.1\tau$ for all emission-line light curves. This half-width value was motivated by previous RM campaigns (Grier et al. 2013). The choice of scaling term makes a strong assumption about the line's responsivity to continuum variations. However, the broad-line responsivity as a function of time and measurable spectroscopic properties is not well understood. Studies that use photoionization models to explore properties of the broad emission lines find that the responsivities are expected to vary between different emission lines (Goad, O'Brien & Gondhalekar 1993; Korista & Goad 2000, 2004). However, these physical properties cannot be easily – or even directly – probed, nor can the responsivity for a particular object and line be predicted a priori. In practice, past RM campaigns have shown variations in the responsivity (i) of the same line between monitoring campaigns (e.g. Gilbert & Peterson 2003; Bentz et al. 2008; Woo 2008; Denney et al. 2010) and (ii) between multiple lines in the same object during the same or temporally similar campaigns (e.g. Cackett et al. 2015; De Rosa et al. 2015). As a result, we acknowledge that our simulation results are sensitive to our simple treatment of the responsivity. Nevertheless, it is neither straight forward, nor clear what the merit may be to apply an alternative, more complex treatment at this time. Additionally, the value of $\Delta C(t)$ implemented in this analysis is only a proxy for the true ionizing flux. As a consequence, the transfer function quoted will be different to the inherent transfer function of the system. We investigate the effects of our choice of transfer function in Section 6.

3.4.3 *Measurement uncertainty*

We created mock light curves with daily cadence, starting ahead of the observing campaign to allow time for the emission-line response (described by the convolution in equation 8) to stabilise before the campaign. We then down-sample the cadence following the OzDES programme design and any extensions described below.

Gaussian errors are added to each mock light curve based on the expected flux uncertainties. The expected flux measurement uncertainty for OzDES RM project is 0.01 mag for photometry and 0.1 mag for spectroscopy. We do not include magnitude-dependent (photon-counting) uncertainties on the line flux measurements as our uncertainties are generally expected to be dominated by our overall absolute flux calibration. The absolute flux calibration of the spectroscopy is performed as follows. (1) The observed F-type stars are matched with an F-type star of equivalent $g - r$ colours (measured from DES photometry) from an existing stellar catalogue. (2) The stellar catalogue spectrum is warped to exactly replicate the colours of the observed F-type star. (3) The observed spectrum is divided by the warped catalogue spectrum. (4) The resulting function is smoothed and represents the sensitivity curve of the observed spectrum. This process is performed separately for the spectra coming from the red and blue arm of AAOmega, and repeated for all the observed F-type stars. (5) The median sensitivity is computed for each season. The sensitivity depends on the observed wavelength and radial position of the fibre, and the median scatter in the sensitivity curve is ∼5 per cent. (6) All AGN spectra are corrected according to the measured sensitivity curve and synthetic g, r, and i magnitudes are calculated. (7) The synthetic magnitudes, and thereby the spectrum, are scaled to the g, r, and i magnitudes measured from the nearest photometry epoch to each spectrum (almost all taken within 4 d) under the assumption that no significant change in flux has occurred. The photometric flux precision in this case is on order of a few per cent. Current OzDES data show that residuals of this calibration process are roughly Gaussian in form, therefore it is reasonable to assume Gaussian uncertainties in our light-curve measurements. We further randomly shifted 10 per cent of the data points without adjusting their error bars to introduce an element of the non-Gaussianities present in real data. The shift was chosen from a Gaussian distribution with a mean value of zero and a standard deviation twice the size of the measurement uncertainty. An example of a simulated light curve is shown in Fig. 6.

The accuracy of the emission-line flux measurement is also sensitive to the method of determining the emission-line flux. The method will differ depending on the quality of the OzDES spectrum. If the SNR in the individual spectrum is sufficient, the spectrum will be modelled to separate the emission line of interest from other AGN, host galaxy, and/or intervening emission and absorption components. Otherwise, the emission-line flux is simply calculated as the integrated flux above an approximate linear continuum fitted between two pseudo-continuum regions near the line. Mg II λ2798 flux measurements are especially affected by Fe II contamination (Vestergaard & Wilkes 2001) and therefore care must be taken when calculating the flux for this line. Additionally, AGN with strong absorption (e.g. broad absorption lines quasars) will be discarded from the final sample as the absorption often alters the emitted line flux.

3.5 Possible survey extensions

We also consider various extensions to the baseline survey to determine if modifications to the programme design can increase the

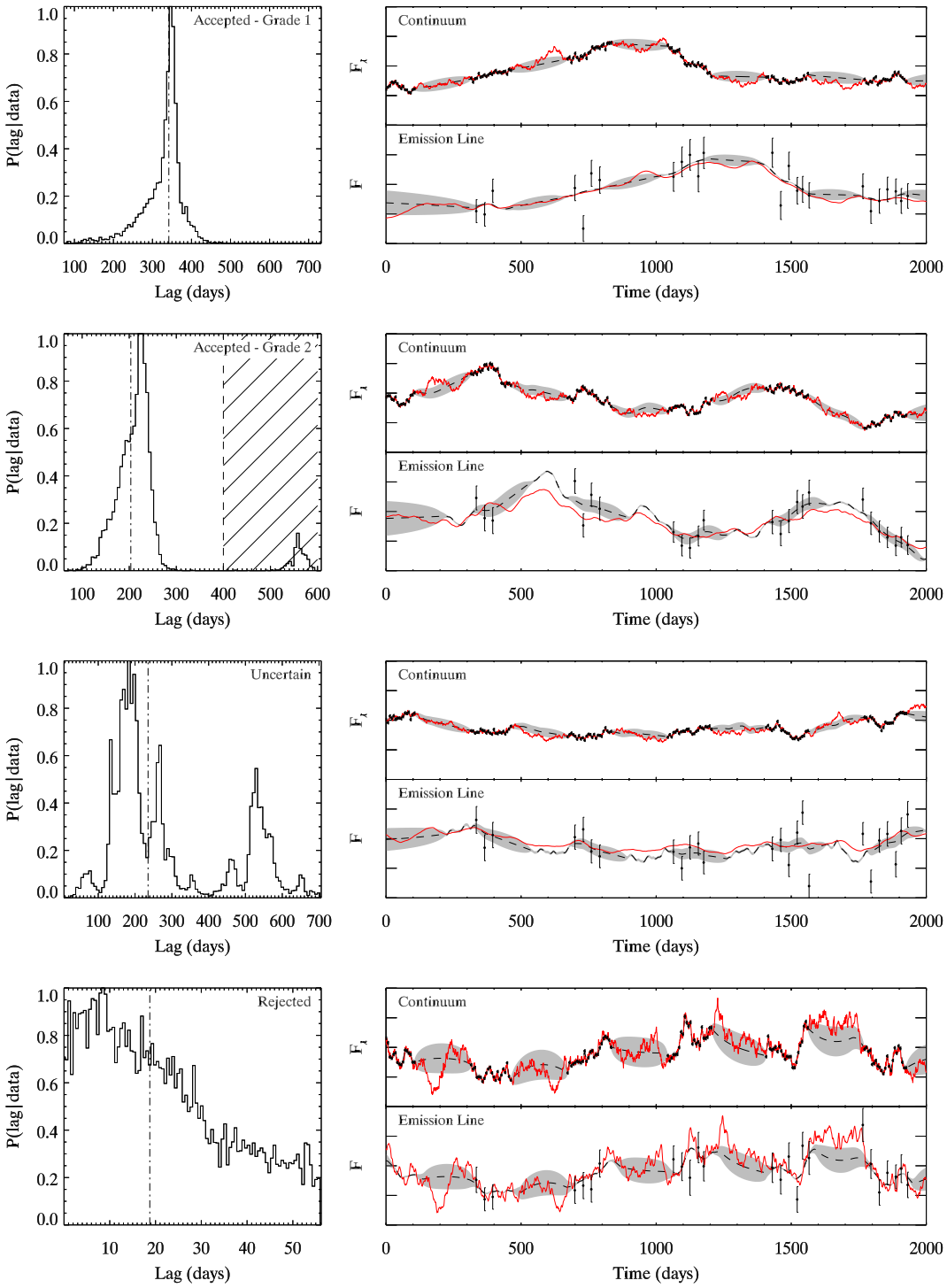


Figure 6. Examples of typical lag posterior distributions (left) for the accepted - grade 1 (top), accepted - grade 2 (middle top), uncertain (middle bottom) and rejected (bottom) classifications. The dot–dashed vertical line represents the true lag value. The hatched region in the accepted - grade 2 case demonstrates the likelihood region excluded from the final lag estimate. The right-hand panels show the input continuum flux density and emission-line flux light curves (black circles with error bars) and true underlying light curve (red) compared to the weighted average of light curves that fit the data well from JAVELIN (black dashed curve), and the corresponding dispersion of these light curves (grey shaded region). The given flux and flux density have arbitrary units.

scientific return. These extensions include (i) decreasing the effective seasonal gap; (ii) increasing the cadence; (iii) improving the data quality; and (iv) extending the total survey duration. While ultimately uncontrollable, we also consider the effect of a reduction in the number of observed epochs due to extreme weather losses.

Seasonal gap (full season/year): we consider two extended spectroscopic observation windows to investigate the effect of the seasonal gap on lag recovery. The longest possible observation window in which any one field can continuously be observed is approximately between May 1 and February 14 for fields E1 and E2 (see

Fig. 2). This will be our first possible extension and will be referred to as 'full season'. As the first two years of observations are completed or underway, we can only apply this extension to the following three years. We also consider a case in which observations are taken over the full year, with no seasonal gaps. This is our second possible extension and will be referred as 'year'. Due to the restricted observability of the DES SN/OzDES fields resulting from their position in the sky, this observational setup is not possible for OzDES, but illustrates the potential for future surveys targeting polar fields. In both cases, it is assumed that photometric measurements are taken weekly and spectroscopic measurements are taken monthly.

Cadence (weekly): we also consider the case where spectroscopic measurements are taken weekly over the baseline observation window of the DES photometry as defined in Section 2. This will test how the spectroscopic measurement frequency affects the recoverability of the lag. Previous RM campaigns have found that high quality and high cadence in the continuum light curve is required for accurate lag recovery as it is the driver of the line light curve. However, it is also important that the line flux measurements are taken with sufficient cadence to map the line response accurately. This possible extension will be referred to as 'weekly'. Again, we can only apply this extension to the last three years of the survey. We should note that this extension increases the number of spectroscopic measurements significantly, which also plays a major role in the recovery of lags (Horne et al. 2004; Shen et al. 2015).

Data quality (goal): we investigate the effects of data quality by changing the spectroscopic measurement uncertainty. We use the baseline sampling rate of OzDES and reduce the spectroscopic measurement uncertainties from 0.1 to 0.03 mag, which corresponds to the optimistic goal for the calibration of the OzDES spectral data. The photometric measurement uncertainty is kept constant at 0.01 mag. This extension will be referred to as 'goal'. We also consider a spectroscopic measurement uncertainty of 0.03 mag combined with the 'year' extension, referred to as 'year+goal'.

Survey length (long): finally we investigate an extension of the survey by two years with sampling and cadence equivalent to the last year of the planned DES/OzDES survey (i.e. seven spectroscopic measurements per additional year). This possible extension will be referred to as 'long' and it should allow the recovery of longer lags and more accurate recovery of lags for a broader AGN population.

Weather: the expected 25 epochs, used in the simulations, already take into account the expected weather loss for the AAT. However, to consider the effects of extreme weather we simulated losing an additional 3–5 spectroscopic epochs randomly over the five-year period.

3.6 Recovering time lags

Traditionally, reverberation lags have been recovered using simple linear interpolation and cross-correlation techniques (Gaskell & Sparke 1986; Edelson & Krolik 1988; White & Peterson 1994), but recently other approaches have been implemented that take into account our existing knowledge of AGN behaviour as a means to optimize, and in some cases, to improve the likelihood of accurate lag recovery. We adopt one such approach, using the program JAVELIN, an updated version of SPEAR (see Zu et al. 2011, for details). Instead of linearly interpolating between data points, JAVELIN uses a DRW to model the AGN continuum light curve, and attempts to fit the emission-line light curve by convolving the continuum light curve with a top hat transfer function.

JAVELIN uses the amoeba minimization method (Press et al. 1992) to recover a model of the continuum light curve (including DRW parameters) and transfer function that best fits the continuum and the line data. JAVELIN has been generally found to be consistent with traditional cross-correlation methods (Zu et al. 2011; Grier et al. 2012; Peterson et al. 2014) and has the advantage of fitting multiple emission-line light curves at once. The main limitation of JAVELIN is that its error estimates assume well-characterized Gaussian noise, so caution must be used when interpreting the parameter uncertainties if these assumptions are violated.

3.6.1 Implementation

We allow JAVELIN to explore the lag range between zero days and three times the input lag, with a maximum allowed lag of 1931 d, corresponding to the separation between the first photometric and last spectroscopic measurement planned for OzDES. Lags longer than 1931 d cannot be constrained by the data; however, we expect very few, if any lags this large in the considered magnitude and redshift range (see Fig. 5). The window of $0 < t < 3\tau$ was found to be sufficient to fully enclose the recovered likelihood distribution of the lags, and there is no evidence for an artificial cut-off in the likelihood distribution (see Fig. 6).

We analysed the five different redshift ranges separately. For the redshift ranges where two emission lines are present, the two lags were fit both individually and simultaneously. For the single line case we used 150 Markov chain Monte Carlo (MCMC) chains with 150 iterations per chain to cover the 'burn-in' period[5] and sample the posterior probability distribution. This is equivalent to 22 500 total 'burn-in' iterations. For the case where both emission-line light curves are fitted simultaneously, we used 200 MCMC chains with 200 iterations per chain. The larger number of iterations account for the additional complexity in the double line fit. For the baseline OzDES setup we analysed 10 realizations of the mock catalogue of 520 AGN. For the investigation of different survey extensions, we only analysed two realizations of the mock catalogue per extension.

3.6.2 Output

For each continuum and emission-line light curve, JAVELIN produces an estimate of the posterior distribution for all the fitted parameters, including the lag. In a substantial number of cases, a clear peak is present and easily identified in the lag posterior distribution, corresponding to the best-fitting lag value. In the remaining cases, multiple peaks of comparable size are present in the distribution or in some rare cases no distinct peak is detected at all.

We devised a simple method to classify each posterior, trained on a manually classified sample of posteriors. In the most basic terms, we smooth the posterior distribution, and identify the dominant peak and any secondary peaks. We classify the quality of the lag based on the existence and relative size of the secondary peaks, which corresponds to the relative probability of the lag being associated with each peak. We classified the output probability distributions into four categories: accepted - grade 1, accepted - grade 2, uncertain, and rejected. Accepted - grade 1 means that no secondary peaks are present in the posterior distribution. Accepted - grade 2 means that a secondary feature is present but the ratio between the

[5] The 'burn-in' period refers to an initial portion of a Markov chain sample that is discarded to minimize the effect of initial values on the posterior inference.

probability of the primary peak and relative probability of the secondary peak is smaller than 15 per cent. The secondary peak may be a bump associated with the main peak. The relative probability of the secondary peak is defined as the difference between the maximum probability value of the secondary peak and either zero or where the secondary connects to the main peak, whichever is the smallest value. This definition is chosen to allow small deviations in the probability distribution of the main peak. Uncertain posteriors have multiple peaks of similar magnitude, while rejected posteriors do not have any clear maxima. The classifications were made without prior knowledge of the true lag. Some examples of typical accepted, uncertain, and rejected lag posterior distributions are shown in Fig. 6. Once classification has been performed, we redefine the lag window boundaries for the accepted - grade 2 cases to exclude the secondary peak where applicable and calculate the credible regions around the primary peak only. This algorithm performs reasonably well compared to human classification, with a 5 per cent misidentification rate. In practice, the real data and lag fits will be subjected to manual inspection which will reduce the number of misidentified lags. The recovered lag used for the rest of this paper is defined as the median lag value from the restricted JAVELIN Monte Carlo chain for accepted lags only, and the lag uncertainty is given by the 16th and 84th percentile range about the primary peak.

3.6.3 *Performance metrics*

To quantify the performance of the OzDES RM campaign and its possible extensions, we define the following three performance metrics: (1) the *recovered fraction*, describing how many lags we recover; (2) σ_Δ, quantifying how accurately the recovered lags are measured; and (3) the number of *misidentified lags* in the recovered fraction.

Recovered fraction: using the classification described in the previous section for accepted, uncertain, and rejected posteriors, we define the recovered fraction as the fraction of accepted - grade 1 and accepted - grade 2 lags. We determine the natural spread in the recovered fraction by performing bootstrap resampling and calculating the recovered fraction for 1000 iterations of the sample. The stated recovered fraction is given as the median recovered fraction from the resampling and the uncertainty is given by the 68th percentile values.

Accuracy (σ_Δ): we next quantify the accuracy of the recovered lags. The accuracy is defined by the logarithmic ratio between the median recovered lag and the true lag,

$$\Delta = \log\left(\frac{\tau^{\rm rec}}{\tau^{\rm real}}\right). \quad (10)$$

Fig. 7 shows that the distribution of Δ values for a large sample of lags is approximately Gaussian, for both the single and two-emission-line simulations, and the skewness of the distribution is minimal. This distribution roughly represents the probability distribution of obtaining a certain accuracy in an individual lag measurement. Since the mean Δ is zero (and the lag recovery is thus unbiased) the underlying width of the distribution, σ_Δ, can be used to characterize the accuracy in the sample, as the lag measurements for the sample tend to be more (less) accurate when σ_Δ is low (high).

Each recovered lag has its own uncertainty, so we cannot simply calculate the unweighted standard deviation. Instead, we find the weighted standard deviation by finding the value of σ_Δ that minimizes the likelihood,

$$\mathcal{L}(\Delta, e_i|\sigma_\Delta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\left(\sigma_\Delta^2 + e_i^2\right)}} \exp\left[\frac{-\Delta_i^2}{2\left(\sigma_\Delta^2 + e_i^2\right)}\right], \quad (11)$$

where $e_i = (\tau^{\rm rec}_{i_{84\,\rm per\ cent}} - \tau^{\rm rec}_{i_{16\,\rm per\ cent}})/2$, is the average measurement uncertainty associated with each lag measurement. To find the uncertainty on σ_Δ we again use bootstrap resampling.

Misidentified lags: a lag is designated as misidentified if it is $3\sigma_\tau$ away from the true lag, where σ_τ is the estimated uncertainty in the lag. This is equivalent to e_i for symmetric uncertainties around the true lag. However, as the uncertainties are asymmetric, we consider the error estimate closest to the true lag value. Note that while we count the number of misidentified lags, we do not remove them from our sample when calculating σ_Δ or any further analysis. This is because we will not know the true lag length when we analyse the real data, and we expect misidentified lags will contaminate the true sample at some level.

3.6.4 *Two-line fitting – simultaneous versus individual fitting*

As previously mentioned, when two emission lines are present in the spectrum we fit the corresponding lags both individually and simultaneously. Simultaneous fitting has the potential to constrain the continuum light curve more tightly, which could enable a better recovery of the two lags. However, we find the accuracy of the recovered lags and the fraction of misidentified lags to be consistent between the two methods, although the simultaneous fits had a lower recovered fraction than the individual fits (Fig. 7). In general, the lag posterior distribution for the simultaneous fitting case showed a peak at the same location as the individual fits, but other secondary peaks were sometimes present, making the lag classification more uncertain. It is likely that this occurs when the lag signal is weak, leading a spurious peak in one lag to amplify an otherwise weak peak in the other lag posterior, leading to multiple peaks. The reverse also occurred, although it was less common, and in this case, the information from the simultaneous fitting nullified the spurious signal. We use the results from the individual line fits for the rest of the analysis.

4 RESULTS

4.1 Basic setup

4.1.1 *Recovered fraction*

We first investigate the recovered fraction in bins of redshift and magnitude. This will help to optimize our target selection choices. The resulting fractions of accepted lags are shown in Figs 8 and 9. There is a clear trend in favour of low-redshift, faint objects.

Since the observer-frame lag also depends on magnitude (through the R–L relationship) and on redshift (due to time dilation), we also investigate how the recovered fraction correlates with the input observed frame lag. Fig. 10 shows the resulting fraction of each lag quality classification (accepted, uncertain, and rejected) as a function of input observed frame lag. There is a steady decline in the recovered fraction with lag length, and a reciprocal increase in the fraction of uncertain and rejected lags. The likelihood of successfully recovering the lag increases when the overall survey length is appreciably longer than the lag length (Horne et al. 2004), because more light curve features can be traced by both the continuum and

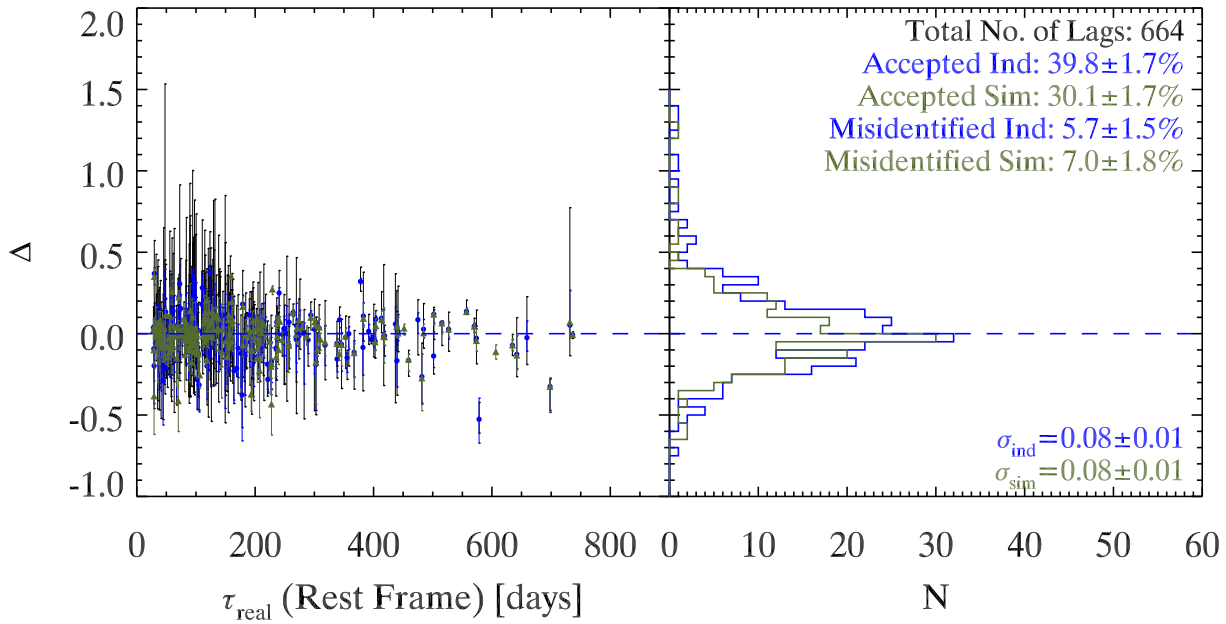


Figure 7. The relative offset of the recovered lag, Δ, for the individually fitted (blue circles) and simultaneously fitted (olive triangles) samples (left). The error bars represent the inner 68 percentile errors. The distribution of Δ is shown in the right-hand panel.

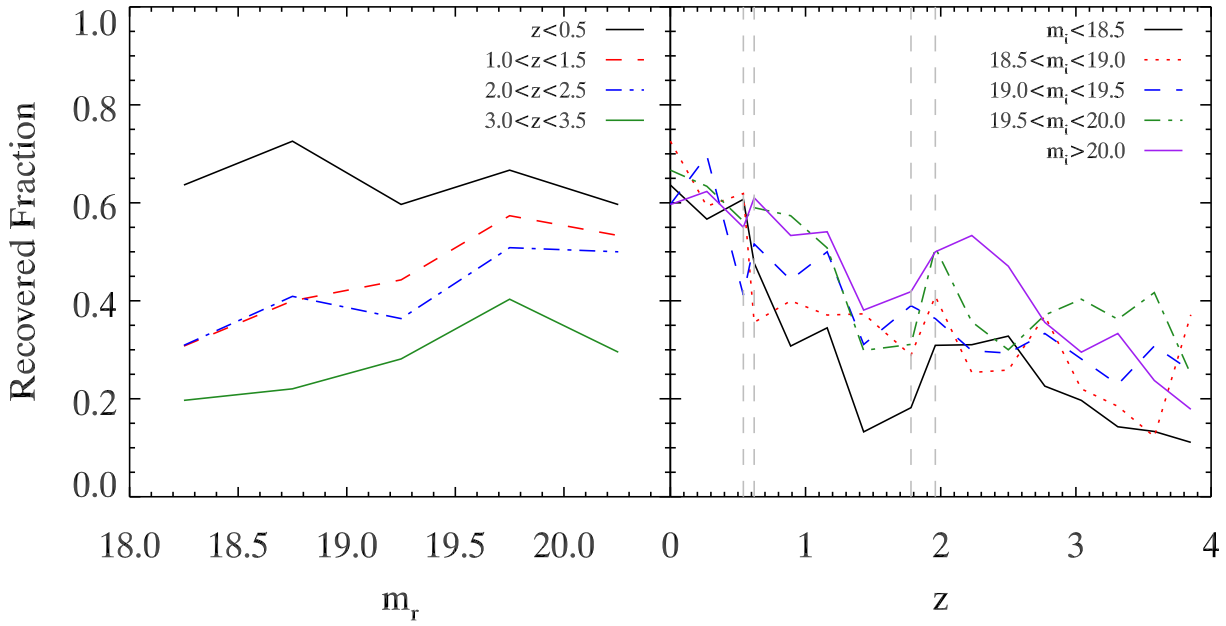


Figure 8. The recovered fraction in each redshift–magnitude bin for the 10 mock catalogues, as a function of magnitude (left) and redshift (right). There is a clear trend both in redshift and magnitude to favour lower redshift, less luminous objects. The dashed vertical lines correspond to the bounds of the different redshift ranges for the different emission lines.

emission-line light curves. However, programme cadence is also important for lag recovery, where sampling the light curves at the 'Nyquist frequency' is required to accurately resolve the lag. Consequently, objects with an expected observer frame lag shorter than the minimum emission-line light curve sampling rate (∼30 d) are less likely to have a reliable lag estimate.

The observed dependence of the recovered fraction on lag length will also be affected by the magnitude dependence of the DRW parameters, specifically $S_F(\infty)$. The $S_F(\infty)$ parameter describes the long term amplitude of the variable continuum light curve component. If $S_F(\infty)$ is small, as is true for more luminous objects, it is less likely that the observed light curve will vary by a significant degree and yield a lag measurement. More accurately, the recovered fraction will be affected by the mean fractional variation of the continuum light curve, $F_{\rm VAR}$, defined as the rms variability amplitude of the continuum magnitude (see Rodríguez-Pascual et al. 1997), rather than the true $S_F(\infty)$ value. By definition, $F_{\rm VAR}$ is closely related to $S_F(\infty)$ as $S_F(\infty) = \sqrt{2}\sigma$, where σ is the long-term standard deviation in the continuum magnitude and $F_{\rm VAR}$ is the fractional standard deviation in continuum flux over a finite observation period. The approximate transformation, $10^{0.4\sigma} \sim 1 + F_{\rm VAR}$, can be used when the observation period is sufficiently long and the

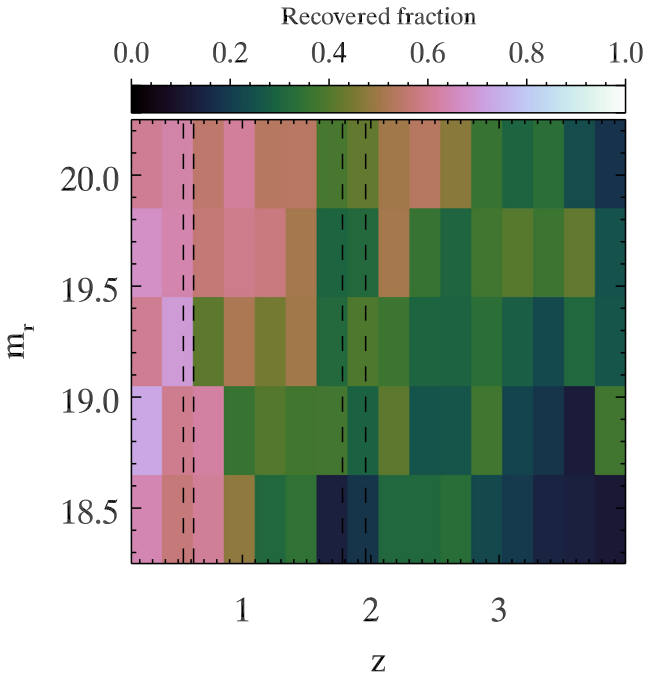


Figure 9. The 2D distribution of the recovered lag fraction in redshift–magnitude bins. This distribution is similar to the distribution of lags (Fig. 5), indicating that the acceptance rate is highly dependent on the observed frame lag.

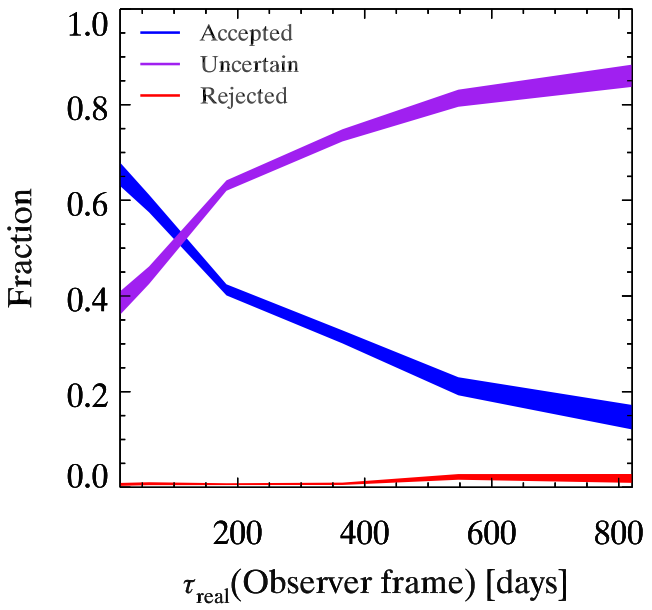


Figure 10. The fraction of accepted (blue), uncertain (purple), and rejected (red) lags as a function of the observed lag.

continuum flux uncertainties are relatively small. AGN with F_{VAR} values lower than the expected measurement uncertainty of OzDES are more likely to have an uncertain lag classification. This is illustrated in Fig. 11, and highlights the importance of minimizing spectroscopic uncertainties to maximize the science return of the RM programme.

In practice, the lag length and light-curve variability are related through the magnitude dependence of the parameters (R–L relationship and equation 5), so the recovered fraction can generally

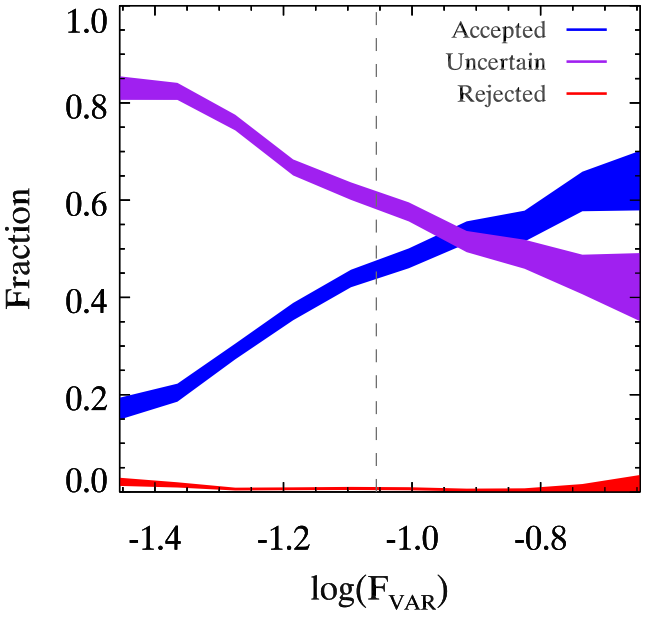


Figure 11. The effect of the rms variability amplitude of the continuum, F_{VAR}, on the fraction of accepted (blue), uncertain (purple), and rejected (red) lag likelihoods. More variable objects are more likely to have a recoverable lag. The dotted line corresponds to the F_{VAR} value equivalent to the assumed emission-line flux uncertainties.

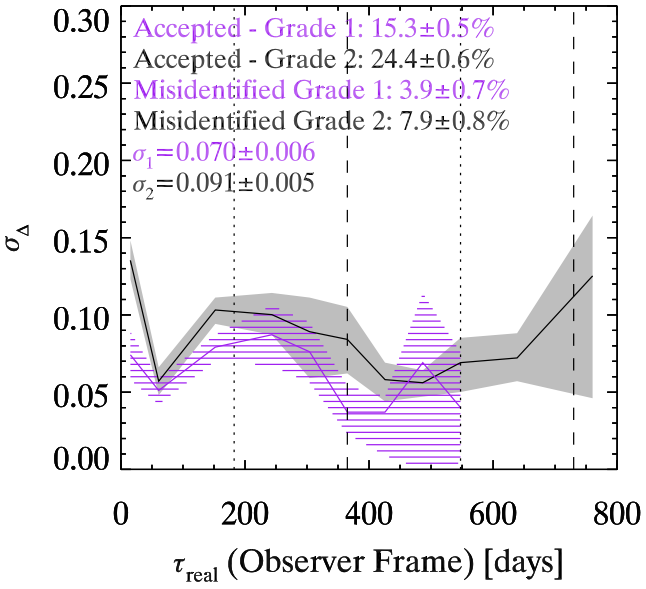


Figure 12. The scatter around the true lag, σ_Δ, as a function of true observed frame lag for the accepted - grade 1 sample (purple hashed) and accepted - grade 2 sample (black solid). The dotted lines correspond to half-year intervals, and the dashed lines correspond to full-year intervals. There are no accepted - grade 1 lags recovered for $\tau_{real} > 1.5$ yr.

be characterized simply by the absolute magnitude and intrinsically brighter objects are less likely to have a lag measurement.

4.1.2 *Accuracy of recovered lags*

Next we look at the accuracy of the recovered lags. Fig. 12 shows how the accuracy changes with observed frame time lag. We divided the sample into accepted - grade 1 and accepted - grade 2.

The accepted - grade 1 lags were more accurate in general, with $\sigma_\Delta = 0.070 \pm 0.006$, compared to $\sigma_\Delta = 0.091 \pm 0.005$ for the grade 2 sample. However, the number of grade 1 lags is almost half the number of grade 2 cases, so excluding the latter would have a significant impact on our final sample size. For both samples we see a slight increase in σ_Δ for $\tau < 30$, and $\tau \sim 0.5$ yr. Lags shorter than 30 d are not accurately recovered due to the limited temporal sampling of the survey. The increase at half a year is likely a consequence of the seasonal gaps in the survey. Lags can only be accurately recovered when there is some overlap between the observed continuum and emission-line light-curve features (e.g. associated rise and drop in light curve). Lags close to $365 \times n/2$ d, where n is an odd number, will have very few observed continuum and emission-line light-curve measurements that can trace out associated light-curve features, leading to larger uncertainties and lower accuracy. Conversely, for lags close to a year, associated light-curve features are well traced by both the observed continuum and emission-line light curve, leading to a well-defined and accurate lag estimate. This effect strongly suggests extending our observing season, as we investigate in the next section.

4.2 Extensions

We next consider how the possible survey extensions affect the lag recovery relative to the baseline survey. The main results from these various extensions are summarized graphically in Figs 13–15, where Fig. 13 shows the distribution of Δ, Fig. 14 shows the recovered fraction, and Fig. 15 shows σ_Δ as a function of the real lag for the different extensions.

Seasonal gap: implementing the full season extension is expected to increase the recovered fraction to ∼50 per cent, which corresponds to a 20 per cent relative increase over the baseline survey. The value of σ_Δ also decreases significantly, especially for lags at $365 \times n/2$ d. This is expected, as the smaller seasonal gaps reduce both the need for interpolation and the range of lags for which common features cannot be traced by both the measured continuum and emission-line light curve. The year extension shows still greater improvements. However, the full season and year extensions only show significant improvement in accuracy for short lags, and no appreciable improvement was seen for long lags. In fact, the year-long extension accuracy was diminished compared to the baseline survey for long lags, as the total duration of the baseline survey is extended beyond five years due the inclusion of the DES science verification data. Without extending the survey length, long lags remain hard to recover because the continuum and emission-line light curves still only have a few common features.

Cadence: the recovered fraction and accuracy of the recovered lags improved significantly with the weekly extension. The finer sampling enables superior recovery of shorter lags, as can be seen in Fig. 14. It is expected that some of this improvement is due to the increased number of epochs. Although, the number of epochs is greater in this scenario than the full season extension the overall improvement in the lag recovery is not as large. This suggests that if additional telescope time is awarded in the next three years of observations, it should be used to extend the observation season rather than to have finer sampling.

Data quality: reducing the measurement uncertainties relative to the baseline significantly increases the recovered fraction and accuracy of the recovered lag. However, improving the measurement uncertainty in the year scenario ('year+goal'), appears to have a negligible effect. In fact, it appears as if a reduction of the emission-line light-curve measurement uncertainty to 0.03 mag has a gain equivalent to carrying out the year survey. Therefore, if we can reduce the uncertainties on the light-curve measurements, additional epochs will be of less consequence. Alternatively, once we reach the limit of reducing the uncertainties, we can improve the scientific outcome with more epochs.

Survey length: the long extension shows an increased recovered fraction and accuracy primarily for the objects with longer lags, as expected. We also see a moderate increase in accuracy for all lags. However, a decrease in recovered fraction and accuracy at half a year is still present due to the seasonal gaps.

Weather: finally, if three to five additional epochs of spectroscopic data were randomly lost over the five years, only a slight drop in recovered fraction and accuracy is expected.

5 PROSPECTS

Maximizing the return is not simply a question of maximizing the number of recovered lags. Next we use the results of the simulations to optimize target selection for measuring black hole masses and constraining the *R*–*L* relationship. Additionally, we fold in the redshift and magnitude distribution of the OzDES target quasars into our predicted results. Fig. 16 shows an example of the recovered lags we expect for the baseline OzDES survey if a final 500 AGN were chosen randomly from the currently observed 989 candidate quasars. The acceptance rate is ∼35 per cent, in this example, and the sample covers a luminosity range of $10^{39} < \lambda L_\lambda(5100\,\text{Å}) < 10^{46}$ with mean $\mu\left\{\log\left[\lambda L_\lambda(5100\,\text{Å})\right]\right\} = 45.0$, and extends to redshift ∼4.

5.1 Black hole mass measurements

The uncertainty in a RM black hole mass estimate is

$$\frac{\sigma_{M_{\rm BH}}}{M_{\rm BH}} = \left[\left(\frac{\sigma_f}{f}\right)^2 + \left(\frac{2\sigma_{\Delta V}}{\Delta V}\right)^2 + \left(\frac{\sigma_R}{R}\right)^2\right]^{1/2}, \quad (12)$$

assuming the errors on the virial factor, f, the line width, ΔV, and BLR radius, R, are independent. Generally, the formal errors reported for $M_{\rm BH}$ only include the uncertainties from R and ΔV and ignore the uncertainty in f (e.g. Bentz et al. 2008; Grier et al. 2012) even though it is generally the largest source of uncertainty in the mass determination. The typical uncertainty in the reverberation masses due to uncertainty in f is ∼0.43 dex (Woo et al. 2010).

The velocity dispersion of gas in the BLR, ΔV, is inferred from the line-of-sight velocity width of the broad emission lines. For RM studies, it is preferred to measure this width in the RMS variance spectrum (henceforth referred to simply as the rms spectrum), as it isolates only those parts of the line that are varying in response to the continuum over the time-scales probed by the campaign. However, this is not always possible because the quality of the rms profile is sensitive to the intrinsic level of variability, the SNR of the data, and the flux calibration uncertainties. Consequently, even if a reliable lag is recoverable from the integrated broad-line flux, the rms profile may be too weak to utilize for a reliable width measurement (e.g. Bentz et al. 2006). Alternatively, interpretation of a width is not always straightforward for complex rms profiles (e.g. Denney et al. 2010). In these cases the velocity width will be measured from the mean campaign spectrum.

Common characterization of this line width include the full width at half-maximum (FWHM), the second moment of the line ($\sigma_{\rm line}$;

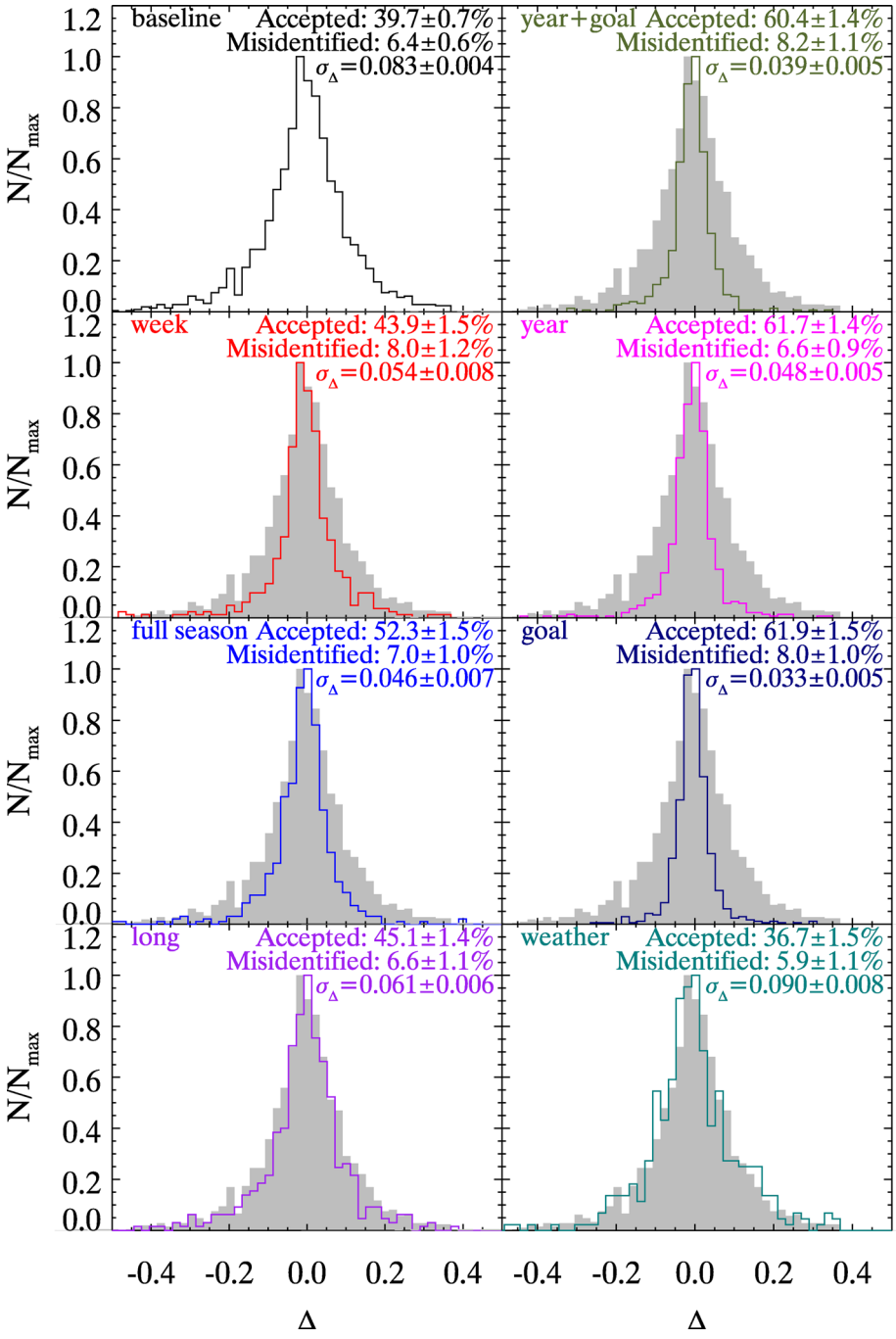


Figure 13. The predicted scatter in the recovered lags and recovered fraction for (from bottom-left to top) the long (two extra years of observations), weather (3–5 epoch of spectroscopic data were lost over the five-year period), full season (nine-month observation season for the last three years of survey), goal (0.03 mag uncertainty in spectroscopic measurement), weekly (weekly cadence of spectroscopic measurements), year (hypothetical five-year survey), and year+goal extensions compared to the baseline OzDES results (grey underlay, and top left).

otherwise known as the line dispersion[6]), and the interpercentile velocity (Whittle 1985). While the precision and accuracy of ΔV depends directly on the SNR of the spectra and the details of the line profile, it also depends on the choice of line-width characterization, since each is susceptible to different systematics (Denney et al. 2009, 2013; Fine et al. 2010; Jensen 2012). We will consider multiple characterizations, though currently most RM studies have preferred the line dispersion as Peterson & Horne (2004) found it follows the best virial relation, and it is the primary characterization used to calibrate the AGN mass scale (Onken et al. 2004; Woo et al. 2010; Park et al. 2012; Grier et al. 2013). The line dispersion has also been found to be a less biased measure of the velocity when measuring Mg II and C IV line widths from

[6] The 'line dispersion' is distinct from 'velocity dispersion' and is defined as, $\sigma_{\rm line}^2 = \langle\lambda\rangle^2 - \langle\lambda^2\rangle = \frac{\int \lambda^2 F(\lambda)\mathrm{d}\lambda}{\int F(\lambda)\mathrm{d}\lambda} - \left(\frac{\int \lambda F(\lambda)\mathrm{d}\lambda}{\int F(\lambda)\mathrm{d}\lambda}\right)^2$.

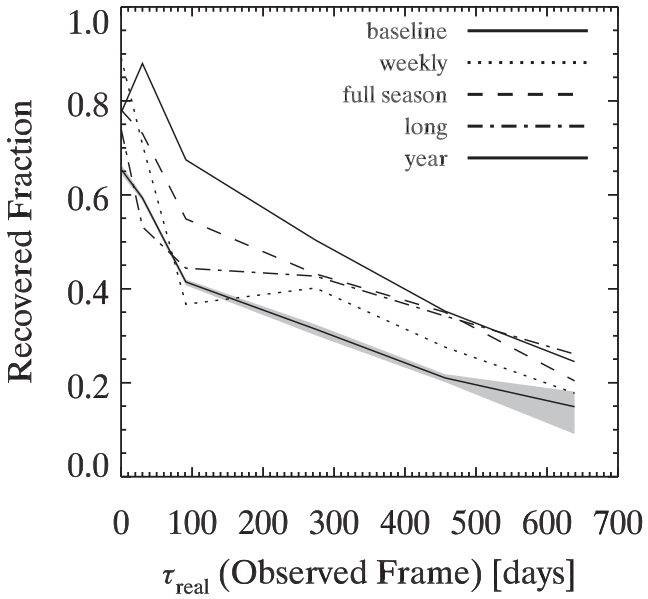


Figure 14. The recovered fraction of the weekly (dotted; weekly cadence of spectroscopic measurements), full season (dashed; nine-month observation season for the last three years of survey), long (dot–dashed; two extra years of observations), and year (triple-dot–dashed; five-year survey with no seasonal gaps and monthly spectroscopic cadence) survey extensions as a function of the observed time lag compared to baseline survey (solid with grey shading; corresponding to the median value and inner 68 percentile values from bootstrap resampling).

single-epoch or mean spectra (Rafiee & Hall 2011; Denney 2012; Denney et al. 2013).

We expect the fractional uncertainty in ΔV, measured from the mean OzDES spectrum to be $\lesssim$5 per cent ($\lesssim$0.04 dex in $M_{\rm BH}$). Compared to the current RM sample, of primarily apparently bright, but low-luminosity objects ($\mu\left\{\log\left[\lambda L_{\lambda}(5100\,\text{Å})\right]\right\} = 43.5$), the level of variability in the OzDES sample is likely to be smaller and the mean spectral SNR lower. It is therefore reasonable to expect the line-width uncertainties from the rms spectra to be appreciably higher, likely on order of 20–30 per cent (0.15–0.20 dex in $M_{\rm BH}$) even for cases with relatively well-defined rms profiles. Due to our lower SNR and larger spectrophotometric errors compared to traditional RM campaigns, we also expect many more cases where we will be unable to recover a width from the rms spectrum, even if we recover a reliable lag.

In cases where traditional rms is too noisy to isolate the broad-line profile, we will first attempt to recover an rms profile from the individual profile model fits to each epoch (see Section 3.4.3), or alternatively use the line widths measured from the mean spectrum. In the latter, it will be important to model and account for absorption and non-variable emission components and use the appropriately calibrated *f*-factor for calculating the black hole mass, as this differs for each line-width characterization and for the mean versus rms spectrum (Collin et al. 2006).

On average, $\sigma_R/R \sim 20$ per cent (0.08 dex in $M_{\rm BH}$) and no significant dependence on redshift is apparent. Therefore, for the baseline OzDES setup we expect a median formal uncertainty in $M_{\rm BH}$ of 0.09 dex and 0.16–0.21 dex in a random sample of OzDES targets using the mean and rms spectrum, respectively (ignoring the uncertainty in *f*).

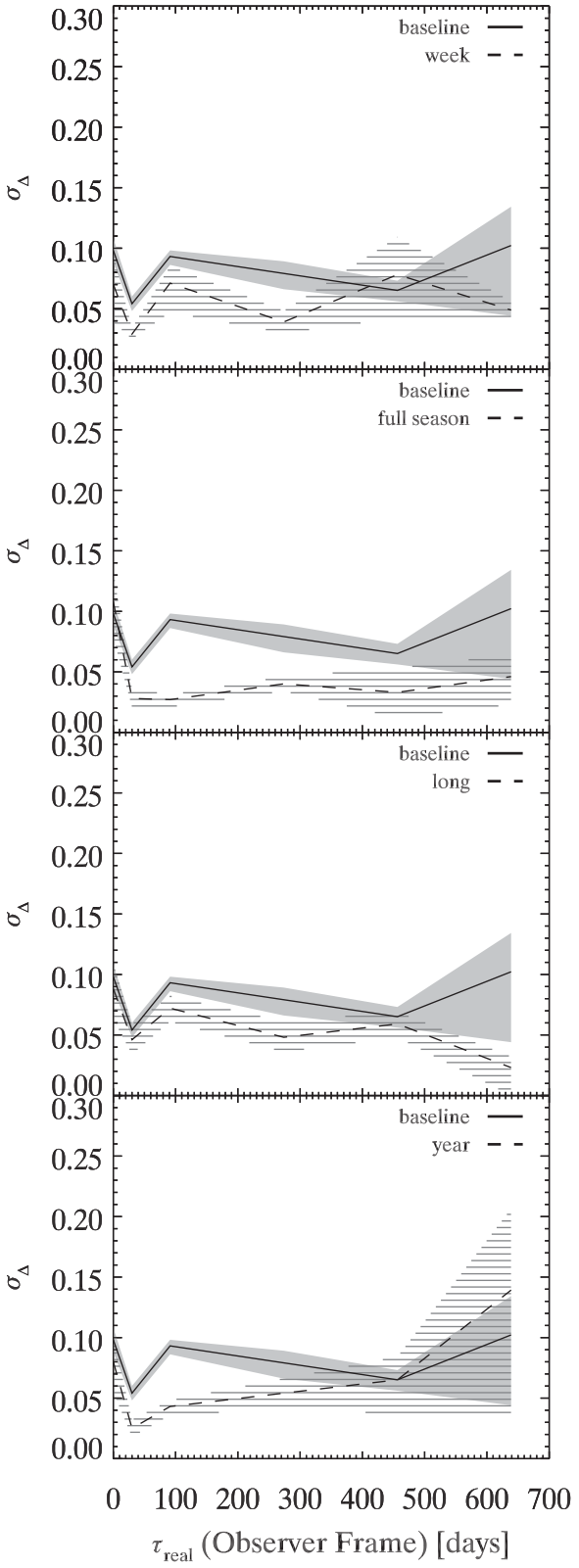


Figure 15. The accuracy σ_Δ of, from top to bottom, the weekly (weekly cadence of spectroscopic measurements), full season (nine-month observation season for the last three years of survey), long (two extra years of observations), and year (five-year survey with no seasonal gaps and monthly spectroscopic cadence) survey extensions as a function of the observed time lag.

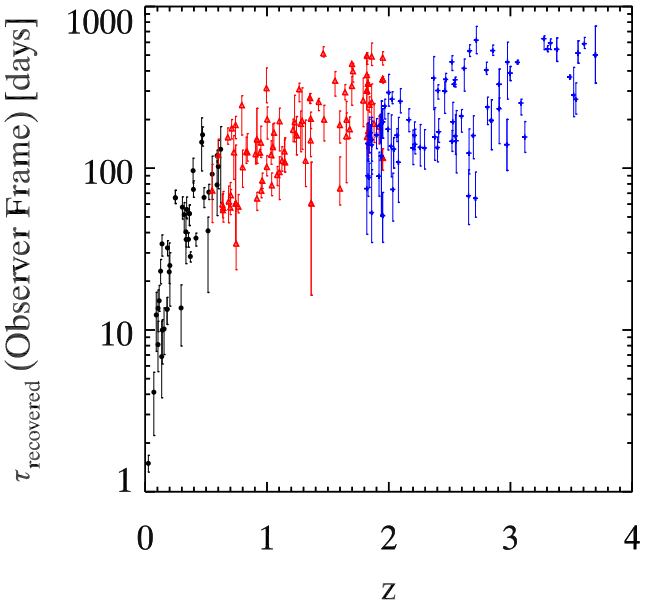


Figure 16. Forecast of lag detections for the OzDES RM campaign. The different colours and symbols correspond to the different emission lines (black circles – H β, red triangles – Mg II, blue crosses – C IV).

5.2 Recovery of the *R–L* relationship

One of the major scientific goals of this survey is to derive the *R–L* relationship for all three emission lines. In this section, we investigate how to optimally select our target quasar sample to recover the most accurate and precise *R–L* relationship. To recover the *R–L* relationship for all three lines we require a substantial calibration sample with lag measurements for two emission lines so that the relationships can be put on the same relative luminosity scale. This is crucial for both black hole mass estimates and any attempts to use quasars as standard candles (King et al. 2014). Therefore we decided to monitor all 89 of the quasars that fall into a redshift range where two lines can be observed simultaneously (i.e. $0.54 < z < 0.62$ [13] and $1.78 < z < 1.96$ [76]).

The optimal strategy for constraining the *R–L* relation is to observe objects over a broad luminosity range. To do so we need to observe both faint low-z objects and bright high-z objects. However, we also want to optimize the number of recovered lags and the accuracy of the lag recovery. For the rest of this section, we determine whether selecting our targets randomly to cover a broad range of properties, or based on their expected acceptance fraction or accuracy, leads to a better estimate of the *R–L* relationship. We separated our 10 AGN mock catalogues into two groups, a training sample and an observed sample. The training sample was used to calculate which redshift–magnitude bins had the highest recovery rates and accuracy. AGN were then selected from the observed sample according to their magnitude and redshift and used to calculate the *R–L* relationship for each emission line. The separation was made in an attempt to avoid any biases in the resulting constraints from particularly favourable or unfavourable bins that do not follow the general magnitude and redshift trends.

The training sample was broken into separate redshift and magnitude bins and the recovered fraction and accuracy, σ_Δ, was calculated for each bin. The resulting distributions were quite noisy, so to remove the influence of spurious bins that may skew our results we fit a low-order polynomial surface to both the recovered fraction and σ_Δ training distributions, using the IDL[7] function SFIT. Due to the significant differences between Mg II and C IV lags, we split the redshift range into a low- ($z < 1.78$) and a high-redshift ($z > 1.96$) group and fit each group separately. The recovered fraction distribution does not appear to have any major structure besides a decline towards bright high-redshift objects, so we simply fit a linear distribution. The distribution of σ_Δ does exhibit several significant features due to the lag dependences found in the previous section, so we fit a third-order polynomial surface to the low-redshift group and a fourth-order polynomial surface to account for the apparent structure in the high-redshift group. The residual distributions of these fits did not show any significant underlying structure.

An observed sample of 500 sources was then either chosen randomly or from specific redshift and magnitude bins according to recovered fraction or accuracy restrictions. The criteria tested were as follows: (a) the recovered-percentage/acceptance was greater than 50 per cent; (b) the recovered-percentage/acceptance was greater than 40 per cent; (c) σ_Δ was less that 0.05; or (d) σ_Δ was less than 0.10.

For each sample we fit an *R–L* relation of the form

$$\log(R/1\,\text{light-year}) = K + \alpha \log(\lambda L_\lambda / 1\,\text{erg s}^{-1}), \quad (13)$$

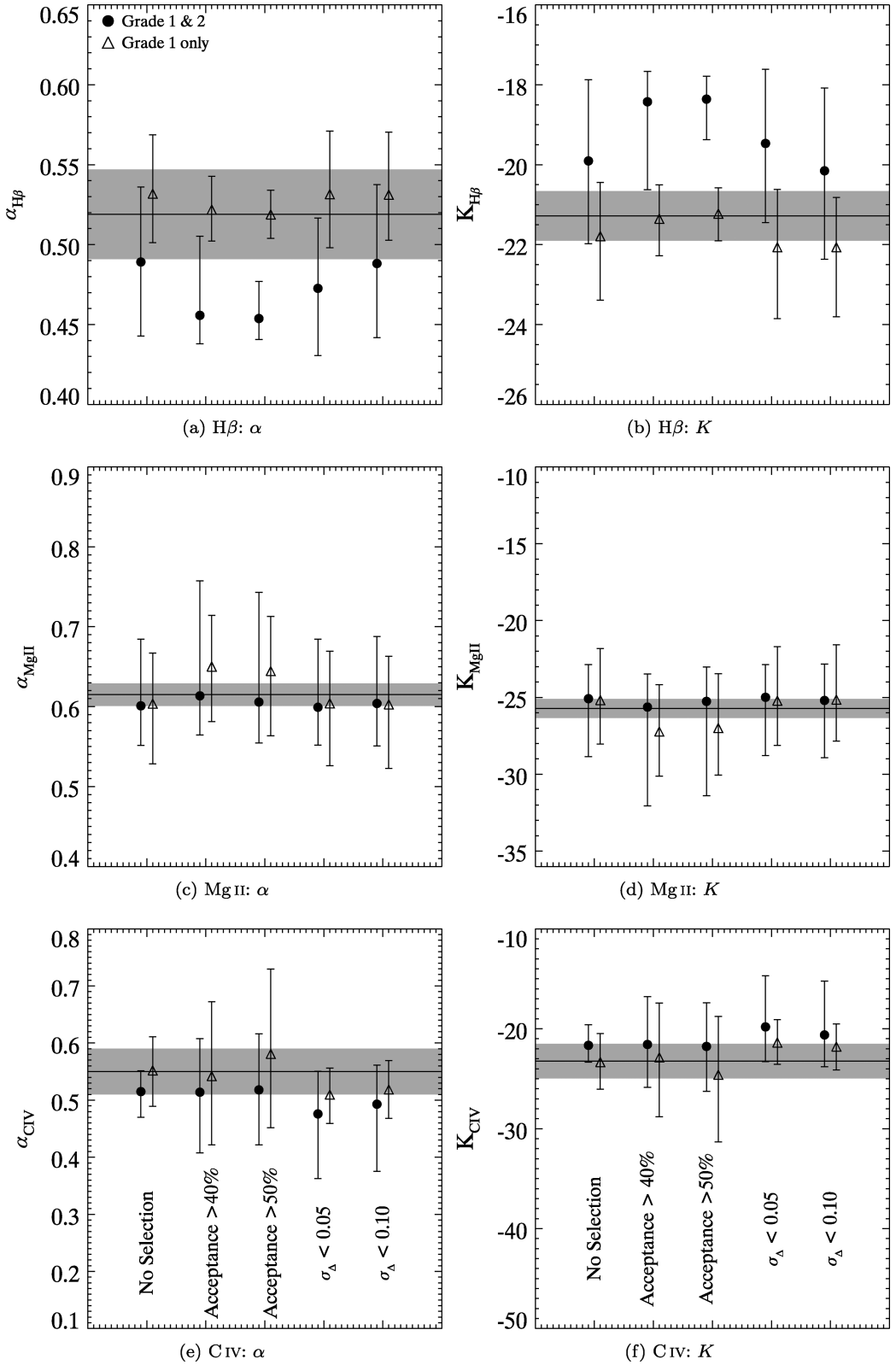
assuming the luminosity of each object is known exactly. The α and K parameters were then calculated for 1000 different possible samples of observed quasars. The resulting median estimates of α and K and their uncertainties are shown in Fig. 17 and Table 3. Following Bentz et al. (2013), we also define the scatter in the sample as the standard deviation of the residuals around the best-fitting relation, and its value is given in Table 3. Table 3 also includes the median observed and expected recovered number of H β, Mg II, and C IV lags for each selection criterion.

For the H β case, there appears to be a systematic bias towards a shallower slope in the recovered *R–L* relationship than expected. This is due to the poor lag accuracy for very short lags, associated with the lowest luminosity, low-redshift objects. We attempt to minimize this effect by restricting our sample to accepted - grade 1 lags only. The resulting α and K values are shown by the triangles in Fig. 17. After this quality cut, the bias is no longer apparent and the recovered values are more tightly constrained. Therefore we should only include the objects with clear RM lag signals. We can also note, that the previous RM campaigns constrain the low-luminosity, low-redshift objects well (Bentz et al. 2013) and their contribution to the *R–L* relation fit has not been considered here.

The Mg II *R–L* relation constraints for both the full accepted sample and accepted - grade 1 only sample were found to be reasonably consistent with the input relationship, although the scatter around the grade 1 best fit was ∼30 per cent smaller than for the full sample. Therefore the Mg II sample is not systematically biased by spurious lag values, but the higher constraining power gained by including more objects is balanced by the increase in the scatter around the relationship.

Like H β though to a lesser extent, the recovered C IV *R–L* parameters also suffered from a bias towards a shallower *R–L* slope with the whole accepted sample, which is resolved by restricting the sample to grade 1. The original bias towards a shallower *R–L* slope is likely due to the short luminosity baseline for C IV *R–L* relationship, which is more sensitive to misidentified lags than the corresponding Mg II *R–L* relation fit. This is despite both lines sharing a similar lag distribution and therefore overall lag accuracy.

[7] Interactive Data Language (Exelis Visual Information Solutions, Boulder, Colorado).



(a) Hβ: α (b) Hβ: K

(c) Mg II: α (d) Mg II: K

(e) C IV: α (f) C IV: K

Figure 17. The median recovered gradient (α: left) and intercept (K; right) of the R–L relation for H β (top), Mg II (middle), and C IV (bottom) for both the full accepted sample (circles) and the accepted - grade 1 sample (triangles) for each sample selection method. The horizontal line and grey shaded regions show the input values for α and K and their current uncertainties.

Table 3. Recovered *R–L* parameters for the five selection criteria.

Selection criteria	Line	K	α	Scatter	Total lags	Accepted lags
None	H β	$-19.9^{+2.0}_{-2.1}$	$0.49^{+0.05}_{-0.05}$	$0.120^{+0.021}_{-0.017}$	80 ± 8	50 ± 6
None	H β - grade 1	$-21.8^{+1.3}_{-1.6}$	$0.53^{+0.04}_{-0.03}$	$0.109^{+0.033}_{-0.023}$	80 ± 8	28 ± 5
None	Mg II	$-25.1^{+2.2}_{-3.8}$	$0.60^{+0.08}_{-0.05}$	$0.122^{+0.018}_{-0.014}$	218 ± 10	88 ± 8
None	Mg II - grade 1	$-25.2^{+3.4}_{-2.9}$	$0.60^{+0.06}_{-0.08}$	$0.092^{+0.016}_{-0.015}$	218 ± 10	34 ± 5
None	C IV	$-21.7^{+2.1}_{-1.7}$	$0.51^{+0.04}_{-0.05}$	$0.126^{+0.014}_{-0.013}$	280 ± 11	96 ± 8
None	C IV - grade 1	$-23.3^{+2.9}_{-2.7}$	$0.55^{+0.06}_{-0.06}$	$0.097^{+0.016}_{-0.015}$	280 ± 11	35 ± 5
Acceptance > 40 per cent	H β	$-18.4^{+0.8}_{-2.2}$	$0.46^{+0.05}_{-0.02}$	$0.125^{+0.013}_{-0.012}$	183 ± 10	115 ± 10
Acceptance > 40 per cent	H β - grade 1	$-21.4^{+0.9}_{-0.9}$	$0.52^{+0.02}_{-0.02}$	$0.114^{+0.020}_{-0.018}$	183 ± 10	65 ± 7
Acceptance > 40 per cent	Mg II	$-25.6^{+2.2}_{-6.4}$	$0.61^{+0.14}_{-0.05}$	$0.127^{+0.014}_{-0.013}$	262 ± 11	124 ± 9
Acceptance > 40 per cent	Mg II - grade 1	$-27.2^{+3.0}_{-2.9}$	$0.65^{+0.06}_{-0.07}$	$0.096^{+0.014}_{-0.014}$	252 ± 11	51 ± 7
Acceptance > 40 per cent	C IV	$-21.6^{+4.8}_{-4.3}$	$0.51^{+0.09}_{-0.11}$	$0.134^{+0.015}_{-0.016}$	134 ± 7	60 ± 7
Acceptance > 40 per cent	C IV - grade 1	$-22.9^{+5.5}_{-5.9}$	$0.54^{+0.13}_{-0.12}$	$0.099^{+0.018}_{-0.017}$	134 ± 7	26 ± 5
Acceptance > 50 per cent	H β	$-18.4^{+0.6}_{-1.0}$	$0.45^{+0.02}_{-0.01}$	$0.121^{+0.010}_{-0.009}$	336 ± 8	220 ± 9
Acceptance > 50 per cent	H β - grade 1	$-21.2^{+0.6}_{-0.7}$	$0.52^{+0.02}_{-0.01}$	$0.116^{+0.015}_{-0.015}$	336 ± 8	112 ± 9
Acceptance > 50 per cent	Mg II	$-25.3^{+2.2}_{-6.1}$	$0.61^{+0.14}_{-0.05}$	$0.105^{+0.019}_{-0.015}$	105 ± 9	86 ± 8
Acceptance > 50 per cent	Mg II - grade 1	$-27.0^{+3.6}_{-3.1}$	$0.64^{+0.07}_{-0.08}$	$0.082^{+0.014}_{-0.012}$	164 ± 9	42 ± 6
Acceptance > 50 per cent	C IV	$-21.8^{+4.4}_{-4.5}$	$0.52^{+0.10}_{-0.10}$	$0.151^{+0.025}_{-0.025}$	70 ± 0	26 ± 4
Acceptance > 50 per cent	C IV - grade 1	$-24.6^{+5.8}_{-6.7}$	$0.58^{+0.15}_{-0.13}$	$0.108^{+0.030}_{-0.033}$	70 ± 0	13 ± 3
$\Delta < 0.05$	H β	$-19.5^{+1.9}_{-2.0}$	$0.47^{+0.04}_{-0.04}$	$0.114^{+0.024}_{-0.018}$	66 ± 7	43 ± 6
$\Delta < 0.05$	H β - grade 1	$-22.1^{+1.5}_{-1.8}$	$0.53^{+0.04}_{-0.03}$	$0.104^{+0.042}_{-0.024}$	66 ± 7	24 ± 5
$\Delta < 0.05$	Mg II	$-25.0^{+2.1}_{-3.8}$	$0.60^{+0.09}_{-0.05}$	$0.120^{+0.015}_{-0.013}$	210 ± 10	85 ± 7
$\Delta < 0.05$	Mg II - grade 1	$-25.2^{+3.5}_{-2.9}$	$0.60^{+0.07}_{-0.08}$	$0.092^{+0.018}_{-0.015}$	210 ± 10	33 ± 6
$\Delta < 0.05$	C IV	$-19.8^{+5.1}_{-3.5}$	$0.48^{+0.07}_{-0.11}$	$0.132^{+0.017}_{-0.015}$	303 ± 10	99 ± 9
$\Delta < 0.05$	C IV - grade 1	$-21.4^{+2.3}_{-2.2}$	$0.51^{+0.05}_{-0.05}$	$0.099^{+0.016}_{-0.014}$	303 ± 10	36 ± 6
$\Delta < 0.10$	H β	$-20.2^{+2.1}_{-2.2}$	$0.49^{+0.05}_{-0.05}$	$0.121^{+0.023}_{-0.017}$	75 ± 7	47 ± 6
$\Delta < 0.10$	H β - grade 1	$-22.1^{+1.2}_{-1.7}$	$0.53^{+0.04}_{-0.03}$	$0.108^{+0.036}_{-0.022}$	75 ± 7	26 ± 5
$\Delta < 0.10$	Mg II	$-25.2^{+2.4}_{-3.7}$	$0.60^{+0.08}_{-0.05}$	$0.124^{+0.019}_{-0.016}$	208 ± 10	84 ± 7
$\Delta < 0.10$	Mg II - grade 1	$-25.1^{+3.6}_{-2.7}$	$0.60^{+0.06}_{-0.08}$	$0.091^{+0.018}_{-0.013}$	208 ± 10	32 ± 5
$\Delta < 0.10$	C IV	$-20.6^{+5.4}_{-3.2}$	$0.49^{+0.07}_{-0.12}$	$0.130^{+0.017}_{-0.013}$	296 ± 10	98 ± 8
$\Delta < 0.10$	C IV - grade 1	$-21.8^{+2.3}_{-2.3}$	$0.52^{+0.05}_{-0.05}$	$0.100^{+0.015}_{-0.016}$	296 ± 10	36 ± 5

The accuracy and precision of the parameter estimates appear to be relatively independent of whether the sample was selected randomly or with an accuracy constraint. However, a sample chosen for high recovered fraction created tighter constraints for H β and poor and possibly misleading constraints in Mg II and C IV. This is primarily driven by the number of lags recovered, although the luminosity baseline of the *R–L* relation fit is also restricted by this choice, which will affect the recovery of the *R–L* constraints. Therefore in order to accurately and precisely recover the *R–L* relationship for all three lines it is preferable to get an even distribution of targets over the total redshift and magnitude distribution rather than maximizing the total number of recovered lags.

5.3 Extensions

We also investigated the improvement in both σ_R/R and the recovered *R–L* parameters associated with the different survey extensions using a random target selection process (equivalent to the 'None' selection process in previous section). The results are summarized in Table 4. The loss of epochs due to weather had a universally detrimental effect on the recovery of $M_{\rm BH}$ and the *R–L* parameters. The precision in the black hole mass measurements was most significantly improved by a reduction in the spectral measurement uncertainty (goal); however, great improvement was also observed in the case where no seasonal gaps are present (year). The 'goal' extension is expected to have a median $M_{\rm BH}$ uncertainty of 0.05 dex, assuming no improvement in the ΔV measurement. Interestingly, the precision in mass from the 'goal' case was found to be superior to the 'year+goal' case, though the difference is not highly significant. The difference is likely due to the slightly longer photometric baseline of the baseline OzDES survey from the DES science verification monitoring in 2012. Note that in the previous section we found the accuracy of the lag measurement to be consistent between the 'goal' and 'year+goal' extensions.

The *R–L* relationship parameter constraints for all lines were most improved from a reduction in the measurement uncertainty (goal; Fig. 18). The precision and accuracy in the H β and C IV *R–L* parameter constraints were also significantly improved in the cases where the seasonal gaps were reduced or removed completely (i.e. year and full season). In the Mg II case, the distribution of constraints for the 'weekly' extension exhibited a large tail towards a shallower *R–L* relationship. This is a consequence of the enhanced lag uncertainty created by the seasonal gap. If only grade 1 lags are used in the construction of the *R–L* relationship, this tail disappears.

Table 4. Recovered *R–L* parameters for the survey extensions.

Extension	$K_{\mathrm{H}\,\beta}$	$\alpha_{\mathrm{H}\,\beta}$	$\mathrm{Scatter}_{\mathrm{H}\,\beta}$	$K_{\mathrm{Mg\,II}}$	$\alpha_{\mathrm{Mg\,II}}$	$\mathrm{Scatter}_{\mathrm{Mg\,II}}$	$K_{\mathrm{C\,IV}}$	$\alpha_{\mathrm{C\,IV}}$	$\mathrm{Scatter}_{\mathrm{C\,IV}}$	σ_R/R
Default	$-19.9^{+2.0}_{-2.1}$	$0.49^{+0.05}_{-0.05}$	$0.120^{+0.021}_{-0.017}$	$-25.1^{+2.2}_{-3.8}$	$0.60^{+0.08}_{-0.05}$	$0.122^{+0.018}_{-0.014}$	$-21.7^{+2.1}_{-1.7}$	$0.51^{+0.04}_{-0.05}$	$0.126^{+0.014}_{-0.013}$	$0.213^{+0.127}_{-0.104}$
Weather	$-22.3^{+4.3}_{-6.3}$	$0.54^{+0.14}_{-0.10}$	$0.147^{+0.051}_{-0.039}$	$-24.8^{+2.0}_{-1.7}$	$0.60^{+0.04}_{-0.05}$	$0.107^{+0.014}_{-0.013}$	$-20.4^{+1.6}_{-1.6}$	$0.49^{+0.03}_{-0.04}$	$0.101^{+0.010}_{-0.009}$	$0.222^{+0.114}_{-0.111}$
Goal	$-21.6^{+0.9}_{-0.7}$	$0.53^{+0.02}_{-0.02}$	$0.069^{+0.021}_{-0.026}$	$-25.5^{+0.5}_{-0.5}$	$0.61^{+0.01}_{-0.01}$	$0.045^{+0.005}_{-0.005}$	$-23.2^{+0.4}_{-0.4}$	$0.55^{+0.01}_{-0.01}$	$0.045^{+0.005}_{-0.005}$	$0.078^{+0.060}_{-0.033}$
Long	$-20.1^{+2.6}_{-2.0}$	$0.49^{+0.05}_{-0.06}$	$0.168^{+0.042}_{-0.040}$	$-23.6^{+0.9}_{-1.0}$	$0.57^{+0.02}_{-0.02}$	$0.110^{+0.015}_{-0.014}$	$-21.6^{+1.2}_{-1.1}$	$0.51^{+0.02}_{-0.03}$	$0.096^{+0.007}_{-0.008}$	$0.174^{+0.076}_{-0.080}$
Weekly	$-21.6^{+1.1}_{-1.1}$	$0.53^{+0.03}_{-0.02}$	$0.094^{+0.025}_{-0.022}$	$-24.8^{+10.3}_{-0.9}$	$0.59^{+0.02}_{-0.23}$	$0.088^{+0.030}_{-0.018}$	$-21.5^{+1.7}_{-1.3}$	$0.51^{+0.03}_{-0.04}$	$0.113^{+0.013}_{-0.012}$	$0.122^{+0.082}_{-0.048}$
Full Season	$-21.9^{+0.8}_{-1.0}$	$0.53^{+0.02}_{-0.02}$	$0.123^{+0.031}_{-0.025}$	$-25.7^{+0.6}_{-0.6}$	$0.62^{+0.01}_{-0.01}$	$0.125^{+0.038}_{-0.056}$	$-24.4^{+0.8}_{-0.9}$	$0.57^{+0.02}_{-0.02}$	$0.084^{+0.012}_{-0.011}$	$0.119^{+0.096}_{-0.051}$
Year	$-20.6^{+0.6}_{-0.6}$	$0.50^{+0.01}_{-0.01}$	$0.058^{+0.008}_{-0.008}$	$-25.2^{+1.5}_{-0.9}$	$0.60^{+0.02}_{-0.03}$	$0.049^{+0.006}_{-0.006}$	$-22.6^{+0.7}_{-0.7}$	$0.54^{+0.01}_{-0.02}$	$0.059^{+0.008}_{-0.007}$	$0.075^{+0.050}_{-0.032}$
Year+goal	$-20.9^{+0.6}_{-0.7}$	$0.51^{+0.02}_{-0.01}$	$0.074^{+0.013}_{-0.012}$	$-25.4^{+0.8}_{-0.8}$	$0.61^{+0.02}_{-0.02}$	$0.063^{+0.045}_{-0.011}$	$-22.7^{+0.6}_{-0.6}$	$0.54^{+0.01}_{-0.01}$	$0.045^{+0.006}_{-0.005}$	$0.076^{+0.065}_{-0.034}$

Note. In this case, the H β sample corresponds to the H β - grade 1 sample.

Additionally, the *R–L* parameter constraints were only marginally improved by extending the survey by an extra two years (long).

6 DISCUSSION AND CONCLUSIONS

The baseline OzDES survey is expected to observe at least 500 AGN for the full five-year period, spanning a redshift range of $0 < z \lesssim 4$ and luminosity range of $10^{39} < \lambda L_{\lambda} < 10^{47}$ erg s^{-1}. If a final 500 AGN are randomly selected from the target quasar catalogue, we expect a lag recovery rate of ∼35–45 per cent (Table 3). This would represent a fourfold increase in the number of measured lags compared to the current sample (Bentz & Katz 2015), and a more than 2-fold increase in redshift range. A higher acceptance rate (∼60 per cent) can be achieved if the final targets are selected based on their expected lag length and their current light curves.

This sample will enable direct M_{BH} measurements over a broad range of quasar properties, environments and black hole masses. It will also help constrain the *R–L* relationship for multiple emission lines and test the robustness of this relationship over a broad AGN population, including an investigation into the recently observed Eddington ratio dependence of the *R–L* relationship discovered by Du et al. (2015). The current RM sample is biased towards local, low-luminosity objects, which are not representative of typical quasars (Richards et al. 2011). This has led to concerns that the existing RM sample may be biased compared to the broader quasar population (e.g. Shen et al. 2008; Richards et al. 2011; Denney 2012). The OzDES sample will probe a similar redshift and luminosity range to the quasar samples in which secondary mass estimate techniques are applied and thus will minimize any potential biases.

The baseline OzDES RM campaign is predicted to accurately recover the *R–L* relationship for all three lines when only the accepted - grade 1 sample is used and a wide distribution of targets are selected (Fig. 17). In general, the spread in the recovered *R–L* parameters is larger than the uncertainties associated with the input relationships. However, the observed OzDES *R–L* relationships will be significantly important as the current Mg II *R–L* relationship is not based on direct Mg II lag measurements (see Trakhtenbrot & Netzer 2012), while OzDES will constrain it directly, and the C IV *R–L* relationship is based only on a small number of objects.

Any of the survey extensions significantly improve the *R–L* relationship parameter constraints for all three lines (Fig. 18 and Table 4). Accurate estimates of all three *R–L* relationships are crucial for single-epoch mass estimates and measuring distances to the highest redshifts. Consequently, there are significant gains from pursuing one or more of the survey extensions we have simulated.

One of the greatest concerns with using AGN as standard candles is whether the *R–L* relationship evolves in redshift. The OzDES sample may enable an investigation into any trends in the *R–L* due to redshift, metallicity, Eddington ratio, and many other properties. If the *R–L* relationship appears to be consistent over the observed sample of quasars, then the OzDES reverberation sample will provide the first physically motivated distance measurement based on a single method from the present day back to redshift four. Unfortunately, the statistical power of the OzDES sample is not expected to be competitive with existing cosmological probes as the predicted uncertainty of the *R–L* will still be too large to rival the precision in current SNe and BAO measurements. However, it has the potential to uncover unexpected expansion behaviour if large deviations to ΛCDM are present at high redshifts (King et al. 2014) and will provide a strong base for future surveys.

The SDSS is currently running a comparable campaign (Shen et al. 2015). They are observing 849 quasars in a 7 deg^2 field of view using the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS) spectrograph. Their sample is flux-limited to $i_{\mathrm{psf}} = 21.7$ mag and includes quasars up to redshift 4.5. Currently they have ∼60 epochs of spectroscopic data taken in 2014 as part of SDSS-III (at ∼4 d cadence), another 12 epochs in 2015 as part of SDSS-IV/eBOSS, and accompanying photometric monitoring taken approximately over the same time period. They expect to recover lags for 10 per cent of their sample out to a redshift of 2, with a possible extension to z∼4 with the inclusion of three years of photometry obtained over 2011–2013 from the Panoramic Survey Telescope and Rapid Response System 1 Survey (Kaiser et al. 2010). The OzDES and SDSS RM campaigns will produce quite complementary measurements. SDSS has finer temporal sampling and a shorter timeline, enabling the recovery of a broad range of faint AGN with shorter lags, while the DES/OzDES sample will be able to more efficiently recover the brighter and higher redshift AGN with longer lags.

6.1 Target selection criteria

Selecting our sample based on the expected accuracy did not significantly change the predicted *R–L* parameter constraints, although, when the sample was chosen for a higher recovered fraction we observe a significant tightening in the H β constraints and degradation of the Mg II and C IV constraints (Fig. 17). This is primarily due to the relative number of lags used to constrain the *R–L* relationships. We find it very advantageous to prioritize targets with multiple lines present in their spectra to calibrate between the H β *R–L* relationship, which is quite tightly constrained by the current RM sample (Bentz et al. 2013), and the *R–L* relationship of the other broad

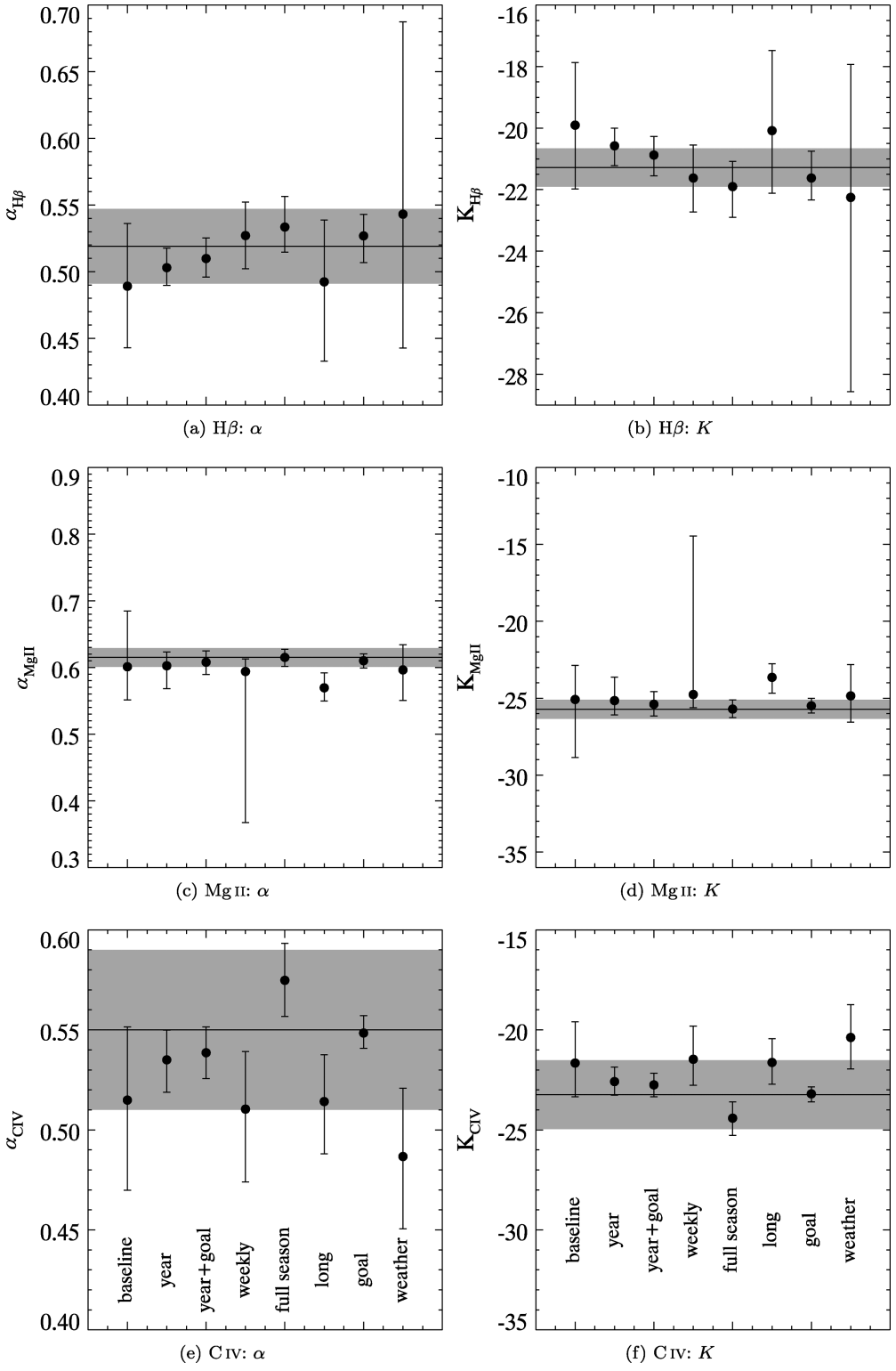


Figure 18. The median recovered gradients (α: left) and intercepts (K; right) of the R–L relation for the various survey extensions. Results are shown for grade 1 H β (top), Mg II (middle) and C IV (bottom) lines. The horizontal line and grey shaded regions show the input values for α and K and their current uncertainty.

lines that currently have little RM data to date. It is also important to cover a diverse range in redshift and magnitude.

The redshift ranges prescribed to each emission line used in this analysis were quite conservative. If they were expanded to allow the largest possible overlap between the lines, while still allowing sensible continuum measurements, the number of AGN with two lines present in their spectrum increases to 216 (i.e. H β–Mg II [13] and Mg II–C IV [203]). However, the quality of these measurements are likely to be lower due to the diminished spectrograph's capabilities at the bluest and reddest wavelengths.

Table 5. Survey extension rankings.

Extension	Resources (add. hours)	Recovered fraction	σ_Δ	σ_R/R	*R–L* accuracy	*R–L* residual scatter	Overall rank
Goal	1	1	1	1	1	1	1
Full season	2	3	2	4	3	3	3
Long	3	4	5	5	4	4	4
Weekly	5	5	4	3	5	5	5
Year	4	2	3	2	2	2	2

Notes. The rankings [1 (best) to 5 (worst)] are given for the required number of additional hours of observation (resources), recovered fraction, accuracy, σ_R/R, and *R–L* parameter recovery. The rank is determined from the median value of each criterion and the final ranking is calculated from the overall sum of the other ranking values.

6.2 Extensions

Of all the survey extensions, improving the measurement uncertainty, through better spectroscopic calibrations, was the most efficient means of improving the overall results. This is because the constraints on each emission-line data point are much stronger, leading to less ambiguity in the lags. The goal measurement uncertainty of 3 per cent is quite optimistic but may be possible to achieve. During the span of the survey so far, major upgrades have been made in both the AAOmega instrument[8] and its pipeline.[9] Another method of reducing these uncertainties is to increase the number of calibrating F-type stars monitored in each field.

In terms of improved scientific results, reducing the measurement uncertainty was closely followed by closing the seasonal gaps, even by three months. By extending the observing season beyond the six-month season of the standard survey we break the 'half-year degeneracy' and allow common variability features to be probed by both continuum and emission-line light curves. However, the improvement in the lag recovery was generally limited to the shorter lags. This result is consistent with previous findings (Horne et al. 2004), namely that the recovery of longer lags requires a longer baseline of observation rather than finer sampling.

Extending the observation campaign by 2 yr (long), only marginally improves the predicted scientific results, despite expectations to the contrary. A longer programme does enable longer lags to be recovered (as seen in Fig. 15), which allows a broader luminosity baseline from which to constrain both the Mg II and C IV *R–L* relationships. However, due to the random selection process we employed and the lower number density of apparently bright objects, this extension had a trivial influence on the *R–L* parameter constraints. It is likely that brighter objects will have relatively higher priority in our target selection (Fig. 1), so in the real survey this improvement may be more substantial. The longest lag, highest redshift objects are also best studied with a long programme like OzDES, and thus there are clear advantages to monitoring as many of those objects as possible in the hope that an extension of the survey will prove possible. The overall ranking of the extensions is shown in Table 5. The rank is determined from the median value of each criterion and the final ranking is calculated from the overall sum of the other ranking values. The criteria tested were, the number of additional hours of observation required (referred to as resources), recovered fraction, accuracy (σ_Δ), precision in lag measurement (σ_τ/τ), and the precision and accuracy of *R–L* parameter recovery. The extensions from best to worst ranking are as follows: Goal, Year, Full season, Long, and Weekly.

[8] http://www.aao.gov.au/science/instruments/current/status
[9] http://www.aao.gov.au/science/software/2dfdr

Losing an additional 3–5 scheduled epochs of spectroscopic data did not significantly affect the recovered fraction and accuracy of the overall mock catalogue sample, but did severely affect both the precision of black hole mass measurements and *R–L* relationship parameters. Therefore it is important to minimize the number of epochs lost over the observation period. One way OzDES is working to minimize potential losses is by working in close collaboration with the 2dFLenS survey, also using the 2dF instrument, to make supplementary observations of the SNe fields when weather restricts OzDES observations. In exchange, OzDES will observe 2dFLenS fields when the DES SNe fields are at high airmass.

6.3 Alternative analysis techniques

OzDES will be one of the longest running RM campaign when it is completed. Nevertheless, the expected number of OzDES spectroscopic epochs is small compared to traditional RM campaigns (e.g. Peterson et al. 2002; Bentz et al. 2009a; Denney et al. 2009; Barth et al. 2011). This leads to the relatively low recovered fraction and low accuracy. However, one way to maximize the output of the OzDES data is to stack the lag signals for multiple objects of similar redshifts and magnitudes ('composite reverberation mapping'; Fine et al. 2012; Brewer & Elliott 2014). Fine et al. (2012) found that by stacking the continuum and emission-line light-curve cross-correlation signals of objects with similar redshifts and magnitudes, and therefore similar expected lags, a mean signal can be recovered even if no lag signal is present in the individual cross-correlations. The large number of OzDES targets makes it a good sample to perform this type of analysis.

We will also search for reverberation signals in other lines in the spectrum beyond H β λ4861, Mg II λ2798, and C IV λ1549. We have focused on these line in this work as they are the three lines traditionally used for single-epoch mass determination, and calibration of the R–L relationship for single-epoch masses is one of the main science drivers of this survey.

6.4 Limitations of survey simulations

There are several limitations to our survey simulation setup. The first is our continuum luminosity determination for the individual objects. We have simply used the existing SDSS template, and have not taken into account host galaxy contamination or extinction. Our choice of quasar template was based on the similar redshift and magnitude range covered by the SDSS sample and the OzDES target sample. The slope of the SED has been found to vary considerably between individual objects and different samples (e.g. Richards et al. 2006) so using a single template is a simplification. However,

it is sufficient for our use as we are investigating the efficiency of the survey for the bulk of the quasar sample.

Not taking into account host galaxy contamination overestimates the bolometric luminosity of the AGN, resulting in reduced sensitivity in our observations, an overestimation of the lag length, and an underestimation of characteristic variability. However, this is only significant for AGN of similar luminosities to their host, and in general, the expected decrease in lag recovery due to lower sensitivity in light-curve variation measurements is counteracted by the increased variability and shortened lag length of lower luminosity quasars. Shen et al. (2015) took this into account in the SDSS RM campaign by assigning a constant host galaxy contribution of 8×10^{43} erg s^{-1} at 5100 Å and negligible contribution at $L_{3000\,\text{Å}}$ and $L_{1350\,\text{Å}}$. Following the approach of Richards et al. (2006, with $L_{\text{bol}}/L_{\text{Edd}} = 0.25$ based on the results of Kollmeier et al. 2006) to estimate the host galaxy contribution of our sample, we find that only low-redshift objects are expected to have a significant host galaxy component and the majority of our targets will not be affected appreciably by neglecting host galaxy light due to their intrinsic luminosities. To confirm this, we tested how the recovered fraction and accuracy of a mock sample were affected by host galaxy contamination and found the results to be consistent with the baseline simulation. However, the inclusion of host galaxy contamination does result in a systematic drop in the expected lag length and again the effect is dominant at low redshifts.

Although host galaxy contamination is not expected to significantly affect the recovered fraction or lag accuracy, not accounting for host galaxy contamination will affect the determination of the optical luminosity for calibration of the H β *R–L* relation parameters. Host galaxy contamination can be estimated, for the observed OzDES data, using a combination of spectral decomposition (Vanden Berk et al. 2006), and high-resolution imaging as performed by Bentz et al. (2009a, 2013).

Ignoring internal extinction also underestimates the bolometric luminosity. However, this is only expected to be a 20 per cent luminosity correction based on extinction estimates for the SDSS DR9 quasar sample (Pâris et al. 2012), which will not cause significant changes to the simulation results.

Another potential limitation of our simulation is our choice of transfer function. A top hat transfer function is a good conservative choice in this type of investigation, as it spreads the lag response more dramatically than a transfer function of a more Gaussian nature. Nevertheless, it could be argued that by choosing a top hat transfer function, when JAVELIN is based on a top hat transfer function, we are biasing our results. Our choice of top hat width may also affect our results. To test both of these issues we consider two alternative transfer functions. First, we considered a Gaussian transfer function, with a width of 0.1τ, motivated by existing RM data (Grier et al. 2013) and the SDSS survey simulations (Shen et al. 2015). The recovered fraction and accuracy were consistent or slightly better than the top hat predictions (Fig. 19). We also investigated a worst-case top hat transfer function scenario, where the top hat width is 2τ instead of 0.2τ. In this case we do see a significant drop in the recovered fraction and accuracy due to the extreme smoothing of the emission-line signal. This case can be considered as the worst-case result for the baseline OzDES survey. We expect negligible effects from changing the width of the transfer function by a factor of 2–3, based on the findings of Shen et al. (2015), who found that lag recovery is not significantly affected by changes in this range.

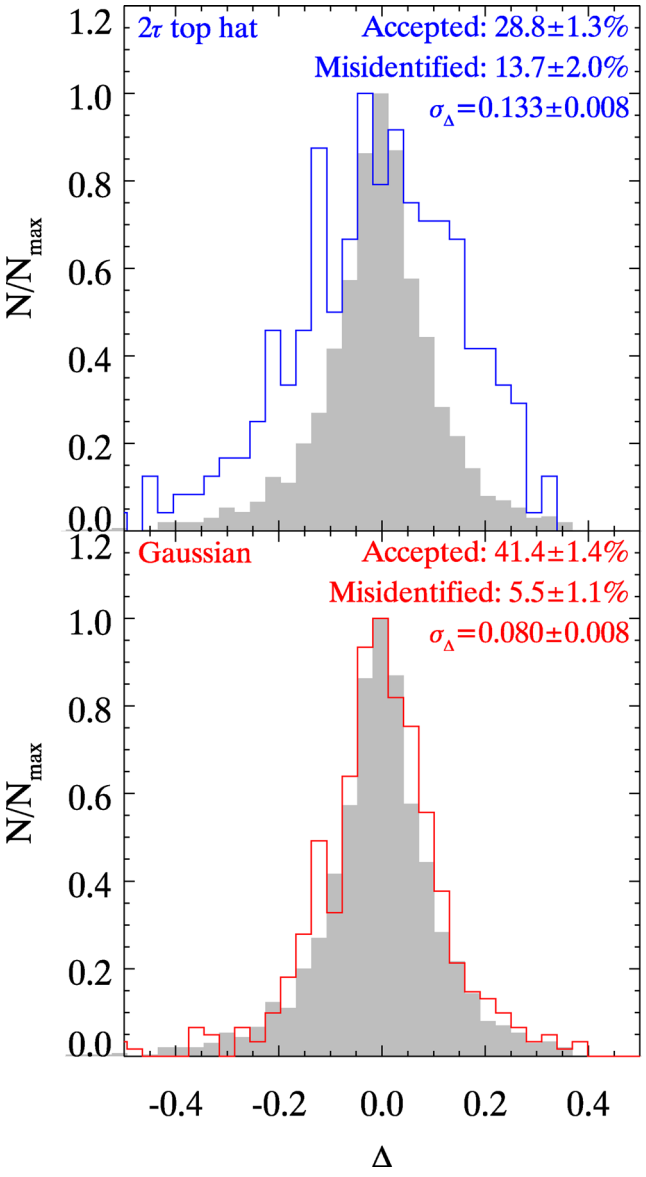


Figure 19. The effects of using a very broad top hat ($w = \tau$) or a Gaussian ($\sigma = 0.1\tau$) transfer function on the simulated results. The results for the baseline 0.1τ top hat transfer function distribution is shown in grey. For comparison, the baseline scenario statistics are as follows: Accepted: 39.7 ± 0.7 per cent; Misidentified: 6.4 ± 0.6 per cent; and $\sigma_\Delta = 0.083 \pm 0.004$.

The choice of top-hat scaling factor will also have a strong effect on the lag recovery. Our choice of unity was based on the analysis of NGC5548 by Zu et al. (2011), which were on the order of 1. However, the authors also found a possible correlation between luminosity and scaling factor that we have neglected in our analysis. The choice of scaling factor directly relates to the responsivity of the line. As mentioned previously, we theoretically expect different emission lines to respond by different degrees to variations in the ionizing continuum (Goad et al. 1993; Korista & Goad 2000, 2004) and this behaviour is witnessed observationally with sometimes contradictory trends. Cackett et al. (2015) recently found that the responsivity of Mg II was low in NGC5548, and displayed virtually no response to the continuum variability over a period of 170 d, despite a near-UV continuum variability amplitude of $F_{\text{VAR}} = 0.33$.

This finding agrees with the observations of Clavel et al. (1991) and Sun et al. (2015) who found a smaller variability amplitude in Mg II flux compared to (i) the other UV lines and (ii) to H β, respectively, though in the case of the latter, possible sample biases preclude whether this effect is intrinsic or not. On the other hand, Woo (2008) found low Mg II responsivity in only one of their objects and reasonable Mg II responsivity in the other four objects. Therefore, defining the responsivity for any line or object appears to be complex. Some objects have also displayed very non-linear responses to continuum variations (e.g. NGC7469, Peterson et al. 2014; NGC5548, De Rosa et al. 2015 during the second half of the campaign; and J080131, Du et al. 2015). Until the mechanisms that drive the response are better understood, we have chosen to adopt a simple parametrization.

6.5 Other considerations

We expect certain sections of the extracted spectra to be suboptimal for flux measurements, in particular the region near the dichroic split of the AAOmega spectrograph at 570 nm and the Fraunhofer A+B absorption bands at 759 and 687 nm. We recommend prioritizing targets whose emission lines avoid these wavelengths by two times the FWHM of the line. This corresponds to approximate redshifts of $z \sim 0.17$ and $z \sim 0.56$ for H β, $z \sim 1.0$ and $z \sim 1.7$ for Mg II and $z \sim 2.7$ and $z \sim 3.9$ for C IV. Avoiding the Fraunhofer bands is of less concern than the dichroic split.

6.6 Summary

We have generated mock catalogues of AGN and corresponding light curves according to the expected OzDES sampling and quasar properties. We attempted to recover the input lag from the simulated light curves to quantify the efficiency and accuracy of the lag recovery. These results were used to predict the expected performance and scientific output of the OzDES RM project and several proposed survey extensions. We expect OzDES to yield lags for ∼35–45 per cent of the monitored quasars. The resulting direct $M_{\rm BH}$ measurements are expected to have formal uncertainties of 0.16–0.21 dex and the baseline OzDES RM campaign will accurately recover the *R–L* relationship parameters for H β, Mg II, and C IV. However, substantial improvements can be gained if we either increase the spectroscopic cadence, extend the survey season, or improve the spectroscopic measurement accuracy of the survey.

ACKNOWLEDGEMENTS

We thank Roberto Assef for useful discussions about the design and the robustness of the survey simulations. ALK would like to acknowledge the support provided by the University of Queensland and Australian Commonwealth Government via the Australian Postgraduate Award. The Dark Cosmology Centre is funded by the Danish National Research Foundation. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. TMD acknowledges the support of the Australian Research Council through Future Fellowship, FT100100595. KDD is supported by an NSF AAPF fellowship awarded under NSF grant AST-1302093. CSK is supported by NSF grant AST-1009756. BMP acknowledges support from NSF grant AST-1008882. MV gratefully acknowledges support from the Danish Council for Independent Research via grant number DFF 4002-00275. Funding for the DES Projects has been provided by the US Department of Energy, the US National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The DES data management system is supported by the National Science Foundation under Grant number AST-1138766. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234, some of which include ERDF funds from the European Union.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

This paper has gone through internal review by the DES collaboration.

REFERENCES

Anderson L. et al., 2012, MNRAS, 427, 3435
Banerji M. et al., 2015, MNRAS, 446, 2523
Barth A. J. et al., 2011, ApJ, 743, L4
Bentz M. C. et al., 2006, ApJ, 651, 775
Bentz M. C. et al., 2008, ApJ, 689, L21
Bentz M. C., Katz S., 2015, PASP, 127, 67
Bentz M. C., Peterson B. M., Netzer H., Pogge R. W., Vestergaard M., 2009a, ApJ, 697, 160
Bentz M. C. et al., 2009b, ApJ, 705, 199
Bentz M. C. et al., 2013, ApJ, 767, 149
Blake C. et al., 2011a, MNRAS, 415, 2892
Blake C. et al., 2011b, MNRAS, 418, 1707
Blandford R. D., McKee C. F., 1982, ApJ, 255, 419
Brewer B. J., Elliott T. M., 2014, MNRAS, 439, L31
Cackett E. M., Gultekin K., Bentz M. C., Fausnaugh M. M., Peterson B. M., Troyer J., Vestergaard M., 2015, preprint (arXiv:1503.02029)
Clavel J. et al., 1991, ApJ, 366, 64
Collier S. J. et al., 1998, ApJ, 500, 162
Collin S., Kawaguchi T., Peterson B. M., Vestergaard M., 2006, A&A, 456, 75
Conley A. et al., 2011, ApJS, 192, 1

Czerny B., Hryniewicz K., Maity I., Schwarzenberg-Czerny A., Życki P. T., Bilicki M., 2013, A&A, 556, A97
Davidson K., 1972, ApJ, 171, 213
De Rosa G. et al., 2015, ApJ, 806, 128
Denney K. D., 2012, ApJ, 759, 44
Denney K. D., Peterson B. M., Dietrich M., Vestergaard M., Bentz M. C., 2009, ApJ, 692, 246
Denney K. D. et al., 2010, ApJ, 721, 715
Denney K. D., Pogge R. W., Assef R. J., Kochanek C. S., Peterson B. M., Vestergaard M., 2013, ApJ, 775, 60
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33
Dietrich M., Kollatschny W., 1995, A&A, 303, 405
Du P. et al., 2014, ApJ, 782, 45
Du P. et al., 2015, ApJ, 806, 22
Edelson R. A., Krolik J. H., 1988, ApJ, 333, 646
Ferrarese L., Ford H., 2005, Space Sci. Rev., 116, 523
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fine S., Croom S. M., Bland-Hawthorn J., Pimbblet K. A., Ross N. P., Schneider D. P., Shanks T., 2010, MNRAS, 409, 591
Fine S. et al., 2012, MNRAS, 427, 2701
Flaugher B. L. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 77350D
Gaskell C. M., Sparke L. S., 1986, ApJ, 305, 175
Gebhardt K. et al., 2000, ApJ, 539, L13
Gilbert K. M., Peterson B. M., 2003, ApJ, 587, 123
Goad M. R., O'Brien P. T., Gondhalekar P. M., 1993, MNRAS, 263, 149
Graham A. W., Erwin P., Caon N., Trujillo I., 2001, ApJ, 563, L11
Grier C. J. et al., 2012, ApJ, 755, 60
Grier C. J. et al., 2013, ApJ, 764, 47
Hinshaw G. et al., 2013, ApJS, 208, 19
Honscheid K., DePoy D. L., for the DES Collaboration 2008, preprint (arXiv:0810.3600)
Hopkins A. M. et al., 2013, MNRAS, 430, 2047
Horne K., Peterson B. M., Collier S. J., Netzer H., 2004, PASP, 116, 465
Jensen J. J., 2012, Master's thesis, Dark Cosmology Centre, University of Copenhagen
Kaiser N. et al., 2010, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Proc. SPIE Conf. Ser. Vol. 7733, Ground-based and Airborne Telescopes III. SPIE, Bellingham, p. 77330E
Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2000, ApJ, 533, 631
Kaspi S., Brandt W. N., Maoz D., Netzer H., Schneider D. P., Shemmer O., 2007, ApJ, 659, 997
Kelly B. C., Shen Y., 2013, ApJ, 764, 45
Kelly B. C., Bechtold J., Siemiginowska A., 2009, ApJ, 698, 895
Kilerci Eser E., Vestergaard M., Peterson B. M., Denney K. D., Bentz M. C., 2015, ApJ, 801, 8
King A., 2003, ApJ, 596, L27
King A., 2005, ApJ, 635, L121
King A. L., Davis T. M., Denney K. D., Vestergaard M., Watson D., 2014, MNRAS, 441, 3454
Kollmeier J. A. et al., 2006, ApJ, 648, 128
Koratkar A. P., Gaskell C. M., 1991, ApJ, 370, L61
Korista K. T., Goad M. R., 2000, ApJ, 536, 284
Korista K. T., Goad M. R., 2004, ApJ, 606, 749
Kormendy J., Gebhardt K., 2001, in Wheeler J. C., Martel H., eds, AIP Conf. Proc. Vol. 586, 20th Texas Symposium on Relativistic Astrophysics. Am. Inst. Phys., New York, p. 363
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Kozłowski S. et al., 2010, ApJ, 708, 927
Krawczyk C. M., Richards G. T., Mehta S. S., Vogeley M. S., Gallagher S. C., Leighly K. M., Ross N. P., Schneider D. P., 2013, ApJS, 206, 4
Krolik J. H., McKee C. F., 1978, ApJS, 37, 459
Laor A., 1998, ApJ, 505, L83
Lewis I. J. et al., 2002, MNRAS, 333, 279
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
MacLeod C. L. et al., 2010, ApJ, 721, 1014
McLure R. J., Jarvis M. J., 2002, MNRAS, 337, 109
McMahon R. G., Banerji M., Gonzalez E., Koposov S. E., Bejar V. J., Lodieu N., Rebolo R., VHS Collaboration, 2013, The Messenger, 154, 35
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Metzroth K. G., Onken C. A., Peterson B. M., 2006, ApJ, 647, 901
Mortlock D. J. et al., 2011, Nature, 474, 616
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
Onken C. A., Ferrarese L., Merritt D., Peterson B. M., Pogge R. W., Vestergaard M., Wandel A., 2004, ApJ, 615, 645
Padmanabhan N., Xu X., Eisenstein D. J., Scalzo R., Cuesta A. J., Mehta K. T., Kazin E., 2012, MNRAS, 427, 2132
Pancoast A., Brewer B. J., Treu T., 2011, ApJ, 730, 139
Pancoast A. et al., 2012, ApJ, 754, 49
Pâris I. et al., 2012, A&A, 548, A66
Park D., Kelly B. C., Woo J.-H., Treu T., 2012, ApJS, 203, 6
Park D., Woo J.-H., Denney K. D., Shin J., 2013, ApJ, 770, 87
Park D., Woo J.-H., Bennert V. N., Treu T., Auger M. W., Malkan M. A., 2015, ApJ, 799, 164
Peterson B. M., 1993, PASP, 105, 247
Peterson B. M., 1999, in Gaskell C. M., Brandt W. N., Dietrich M., Dultzin-Hacyan D., Eracleous M., eds, ASP Conf. Ser. Vol. 175, Structure and Kinematics of Quasar Broad Line Regions. Astron. Soc. Pac., San Francisco, p. 49
Peterson B. M., Horne K., 2004, Astron. Nachr., 325, 248
Peterson B. M., Wanders I., Bertram R., Hunley J. F., Pogge R. W., Wagner R. M., 1998, ApJ, 501, 82
Peterson B. M. et al., 1999, ApJ, 510, 659
Peterson B. M. et al., 2002, ApJ, 581, 197
Peterson B. M. et al., 2004, ApJ, 613, 682
Peterson B. M. et al., 2014, ApJ, 795, 149
Planck Collaboration XVI, 2014, A&A, 571, A16
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd edn. Cambridge Univ. Press, Cambridge
Rafiee A., Hall P. B., 2011, ApJS, 194, 42
Rafter S. E., Kaspi S., Behar E., Kollatschny W., Zetzl M., 2011, ApJ, 741, 66
Rafter S. E., Kaspi S., Chelouche D., Sabach E., Karl D., Behar E., 2013, ApJ, 773, 24
Reichert G. A. et al., 1994, ApJ, 425, 582
Richards G. T. et al., 2006, AJ, 131, 2766
Richards G. T. et al., 2011, AJ, 141, 167
Richstone D., 1998, in Sofue Y., ed., Proc. IAU Symp. 184, The Central Regions of the Galaxy and Galaxies. Kluwer, Dordrecht, p. 451
Robinson A., 1994, in Gondhalekar P. M., Horne K., Peterson B. M., eds, ASP Conf. Ser. Vol. 69, Reverberation Mapping of the Broad-Line Region in Active Galactic Nuclei. Astron. Soc. Pac., San Francisco, p. 147
Rodríguez-Pascual P. M. et al., 1997, ApJS, 110, 9
Saunders W. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 389
Schulze A., Wisotzki L., 2010, A&A, 516, A87
Shen Y., Liu X., 2012, ApJ, 753, 125
Shen Y., Greene J. E., Strauss M. A., Richards G. T., Schneider D. P., 2008, ApJ, 680, 169
Shen Y. et al., 2011, ApJS, 194, 45
Shen Y. et al., 2015, ApJS, 216, 4
Skielboe A., Pancoast A., Treu T., Park D., Barth A. J., Bentz M. C., 2015, preprint (arXiv:1502.02031)
Smith G. A. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 410
Stirpe G. M. et al., 1994, ApJ, 425, 609
Sun M. et al., 2015, preprint (arXiv:1506.07886)
Sutherland W. et al., 2015, A&A, 575, A25

Taylor K., Cannon R. D., Watson F. G., 1997, in Ardeberg A. L., ed., Proc. SPIE Conf. Ser. Vol. 2871, Optical Telescopes of Today and Tomorrow. SPIE, Bellingham, p. 145
Trakhtenbrot B., Netzer H., 2012, MNRAS, 427, 3081
Trevese D., Perna M., Vagnetti F., Saturni F. G., Dadina M., 2014, ApJ, 795, 164
Trump J. R. et al., 2011, ApJ, 733, 60
Vanden Berk D. E. et al., 2001, AJ, 122, 549
Vanden Berk D., Yip C., Connolly A., Jester S., Stoughton C., 2004, in Richards G. T., Hall P. B., eds, ASP Conf. Ser. Vol. 311, AGN Physics with the Sloan Digital Sky Survey. Astron. Soc. Pac., San Francisco, p. 21
Vanden Berk D. E. et al., 2006, AJ, 131, 84
Vestergaard M., 2002, ApJ, 571, 733
Vestergaard M., 2004, ApJ, 601, 676
Vestergaard M., Osmer P. S., 2009, ApJ, 699, 800
Vestergaard M., Peterson B. M., 2006, ApJ, 641, 689
Vestergaard M., Wilkes B. J., 2001, ApJS, 134, 1
Vestergaard M., Fan X., Tremonti C. A., Osmer P. S., Richards G. T., 2008, ApJ, 674, L1
Wandel A., Peterson B. M., Malkan M. A., 1999, ApJ, 526, 579
Wanders I. et al., 1997, ApJS, 113, 69
Warren S. J., Hewett P. C., Foltz C. B., 2000, MNRAS, 312, 827
Watson D., Denney K. D., Vestergaard M., Davis T. M., 2011, ApJ, 740, L49
White R. J., Peterson B. M., 1994, PASP, 106, 879
Whittle M., 1985, MNRAS, 213, 1
Willott C. J. et al., 2010, AJ, 140, 546
Woo J.-H., 2008, AJ, 135, 1849
Woo J.-H. et al., 2010, ApJ, 716, 269
Wright E. L. et al., 2010, AJ, 140, 1868
Yuan F. et al., 2015, MNRAS, 452, 3047
Zu Y., Kochanek C. S., Peterson B. M., 2011, ApJ, 735, 80
Zu Y., Kochanek C. S., Kozłowski S., Udalski A., 2013, ApJ, 765, 106

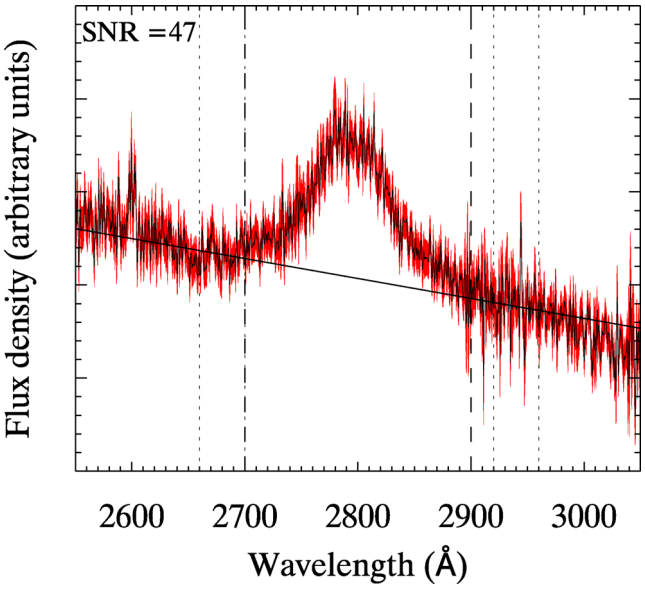


Figure A1. An example of the line and continuum decomposition implemented in the SNR measurement process. The diagrams shows only the Mg II λ2798 emission line. The spectrum is shown by the black solid curve. The associated flux density uncertainty, $\sqrt{\mathrm{Var}(F_\mathrm{L})}$, is shown by the surrounding red shaded region. The line summation limits and associated continuum regions are shown by the dashed and dotted vertical lines, respectively. The linear continuum fit is shown by the solid line and the resulting SNR in the emission-line flux measurement for this example is 47.

APPENDIX A: QUALITY OF OZDES SPECTRA

To quantify the quality of the OzDES spectra we estimated the SNR of the emission-line flux in the rest-frame spectrum, as follows.

(i) For each line, the continuum flux density (S^C) is defined by a linear function of wavelength anchored by the median pixel value in the upper and lower continuum regions, defined in Table A1.

(ii) The line flux, F_L, is measured as the integrated flux above the linear continuum flux within the specified line integration wavelength limits such that

$$F_\mathrm{L} = \sum_{i_1}^{i_2} \left[S_i - S_i^\mathrm{C}\right], \tag{A1}$$

where S_i is the uncalibrated flux density value at pixel i, and i_1 and i_2 are the pixel values that correspond to the defined minimum and maximum emission-line wavelength range.

(iii) The associated variance in the line flux, $\sigma^2(F_\mathrm{L})$, is measured as the integrated variance within the emission-line wavelength region. The SNR is then calculated as

$$\mathrm{SNR} = \frac{F_\mathrm{L}}{\sqrt{\sigma^2(F_\mathrm{L})}}. \tag{A2}$$

Table A1. Line summation limits.

Emission line	Line integration limits (Å)	Continuum window Lower (Å)	Continuum window Upper (Å)
H β λ4861	4810–4940	4770–4800	5100–5130
Mg II λ2798[a]	2700–2900	2660–2700	2920–2960
C IV λ1549	1470–1620	1440–1470	1700–1730

[a]The regions around Mg II are severely affected by Fe II contamination so this choice of pseudo-continuum wavelength range should be used critically (Vestergaard & Wilkes 2001).

An example of a spectrum and the associated line summation limits and continuum fits is given in Fig. A1. The SNR for a single object can vary markedly between lines and epochs due to changes in observation conditions and individual line strength. Therefore, we allocate an SNR value for each quasar based on the median SNR of the best measured line in the spectra. Fig. A2 shows the magnitude and redshift distribution of quasars with different SNR value cut-offs. There appears to be no marked difference between the redshift distributions of high- and low-SNR objects; however, a significant shift in magnitude is observed such that higher SNR objects tend to be brighter, as expected.

These signal-to-noise measurements are preliminary, and only serve to quickly quantify the expected distribution of emission line measurement quality within the OzDES sample. Line strength measurements will improve with the more thorough methods we will utilize for the final RM analysis, including decomposition of the spectra to remove contamination.

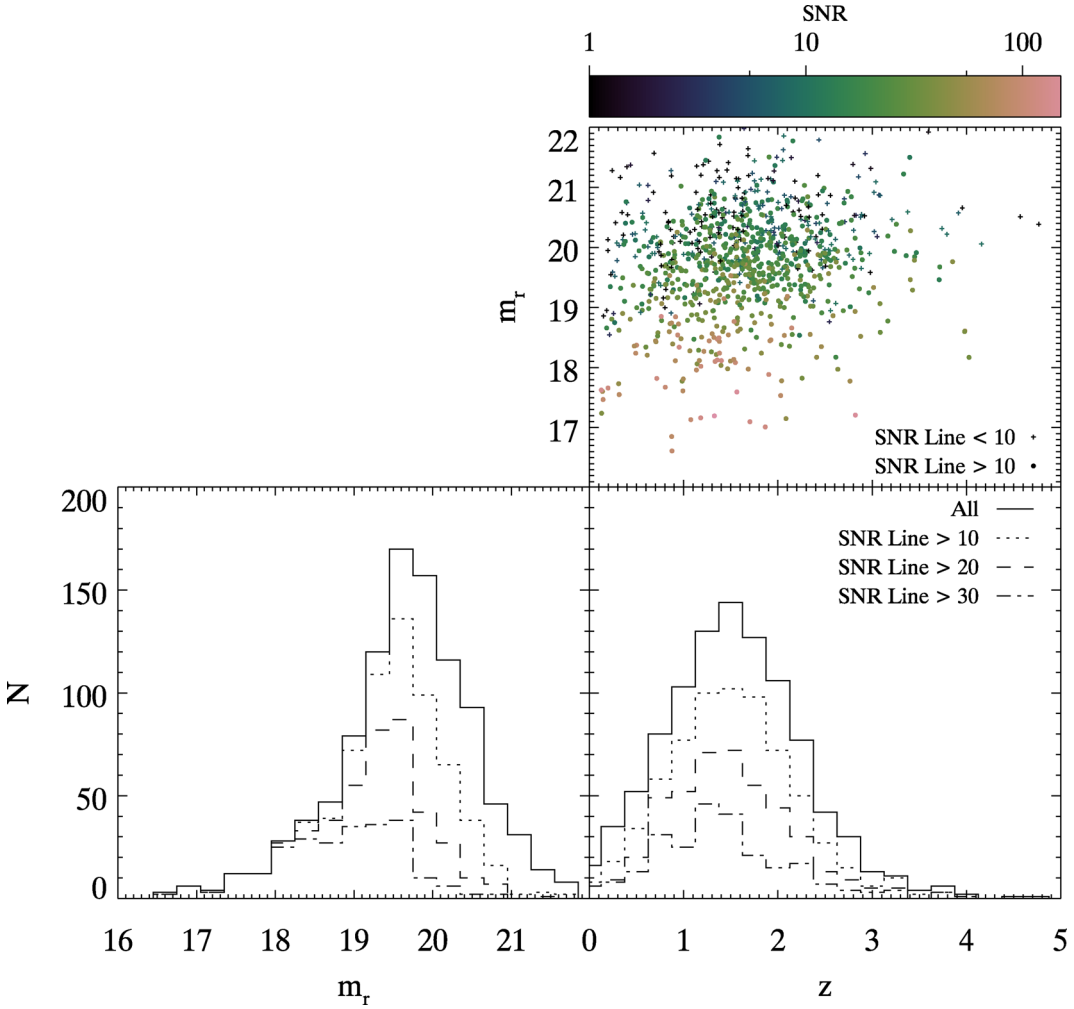


Figure A2. The top panel shows the distribution in characteristic emission-line flux measurement signal-to-noise (SNR) values for each object, the dots represent SNR $>$ 10 and crosses represent SNR $<$ 10. To achieve the baseline emission-line uncertainty of 0.1 mag a bare minimum of SNR $>$ 10 is required. The histograms show the magnitude distribution (left) and redshift distribution (right) of the whole sample (solid) compared to objects with SNR values greater that 10 (dotted), 20 (dashed), or 30 (dot–dashed).

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