

OPTIMAL MEASUREMENTS OF REDSHIFTS USING THE WEIGHTED CROSS-CORRELATION

DANIEL D. KELSON, PAUL MARTINI¹, & J. S. MULCHAEY

Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101

kelson@ociw.edu, pmartini@cfa.harvard.edu, mulchae@ociw.edu

Draft version of May 5, 2005

ABSTRACT

A large component of astronomy involves the measurement of redshifts using absorption line spectroscopy. Typically such data have non-uniform sources of noise and other systematic defects not easily dealt with when one employs Fourier-based techniques because such methods explicitly weight the data uniformly. Here we develop a method for the measurement of redshifts using the cross-correlation in the Real domain, in which one is free to employ non-uniform weighting. The implementation we describe in this paper allows for the arbitrary exclusion of bad data, and weights each remaining pixel by the inverse of the variance. This prescription for weighting the pixels has the advantage that the units of the cross-correlation are exactly half that of χ^2 . Thus, the topology of the peak of the weighted cross-correlation is directly related to the confidence limits on the measured redshifts. The validity of the redshifts and formal errors derived with this method are tested using simulations of galaxy spectra with a broad range of signal-to-noise ratios. These simulations also include tests of the effects of template mismatch. Overall, template mismatch is only significant when the data have high signal-to-noise ratios, and in such cases the systematic error due to mismatch is minimized when one chooses the template that minimizes the error in the redshift. This method is applied to a subset of the data taken in our survey of galaxies in the field of Abell 2550. Based on 36 redshifts, we find that galaxies in the field of Abell 2550 are actually distributed in four distinct groups of galaxies between redshifts $0.1 \lesssim z \lesssim 0.2$ rather than in a monolithic, rich cluster. While the weighted cross-correlation has been discussed in the context of extragalactic redshift surveys, this method may also prove useful in measuring the radial velocity of stars and other astronomical objects.

Subject headings: methods: data analysis — techniques: spectroscopic — galaxies: clusters: individual (Abell 2550)

1. INTRODUCTION

Measuring the radial velocities of astronomical objects has been a key component of modern astrophysics for almost 150 years (for a compendium of early history see Griffin 1967). Currently redshifts are most often measured by cross-correlating the spectrum of one's target with that of a template (Simkin 1974; Tonry & Davis 1979; Kurtz & Mink 1998, and others). Traditionally such cross-correlations have been performed in the Fourier domain, as scarce computational resources can be most efficiently used by employing Fast Fourier Transforms (Cooley & Tukey 1965; Press et al. 1986). Prior to performing the cross-correlation, both target and template spectra are filtered and normalized, producing a cross-correlation for which there no longer exists an obvious relationship between the properties of the cross-correlation and the expected statistics of noise in the data of the target. Without a simple connection between the properties of the cross-correlation and the statistics of noise in the data, Tonry & Davis (1979) introduced the formalism for extracting formal errors and confidences in the estimated locations of cross-correlation peaks (also see Heavens 1993). This formalism has served the astronomical community very well and is by far the most commonly employed technique for the measurement of redshifts.

One alternative/modification was introduced by Glazebrook, Offer & Deeley (1998) in which galaxy spectra

are represented as a superposition of orthogonal template spectra rather than by a single template spectrum. The goal in their work was to minimize the potentially large errors arising from template mismatch. Glazebrook, Offer & Deeley (1998), attempted to recast the discussion of cross-correlation in the context of χ^2 statistics, but their method is still inherently a refinement of the cross-correlation method performed in the Fourier domain.

Fourier-based techniques, however, suffer from an important limitation: they treat each pixel with equal weight. Thus, localized sources of noise, such as from bright night-sky emission lines, have greater weight in the cross-correlation than one would choose *a priori*. Furthermore, the presence of emission lines leads to peculiar complications and epicycles in the published procedures (e.g. Kurtz & Mink 1998). Even with the method outlined in Glazebrook, Offer & Deeley (1998), the mismatch of emission line strengths between the template(s) and the galaxy spectrum can lead to problems. Spurious results may best be avoided if pixels are weighted by the signal-to-noise ratio, or ignored altogether if sufficiently contaminated by night-sky emission or other sources of deviance. Techniques using Fast Fourier Transforms simply do not offer the flexibility for such an optimal/weighting approach.

Fortunately, the abundance of computational resources in modern astronomy no longer necessitates the use of Fourier techniques for redshift measurements. In that spirit, we introduce a weighted cross-correlation scheme,

¹ present address: Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

taking full advantage of the benefits of working in the Real domain. The basic procedure is outlined in §2, followed by verification of the estimates of the formal errors using simulations in §3. In §4 we illustrate the application of this technique to spectroscopic observations of a sample of galaxies in the field of Abell 2550. The key points are summarized in the last section.

2. THE BASIC PROCEDURE

The cross-correlation, \mathcal{C} , of a galaxy spectrum, G , and a template spectrum, S , is denoted by $\mathcal{C}(z)$:

$$\mathcal{C} = \mathcal{L}(S) \circ \mathcal{L}(G) \quad (1)$$

where $\mathcal{L}(X)$ is an operation that eliminates low-frequency structure in X , leaving a spectrum of mean zero. This template spectrum, S , has been rebinned to the logarithmic wavelength binning of G , as is the case for the Fourier techniques mentioned above. In Tonry & Davis (1979) $\mathcal{L}(X)$ are normalized by the rms of the filtered spectrum. The discrete form of the cross-correlation is an explicit summation over n pixels, each labeled i :

$$\mathcal{C}(z) = \sum_{i=1}^n \mathcal{L}(T)_i \mathcal{L}(G)_i \quad (2)$$

where now $T = \mathcal{V}(S, z)$ is the operation that redshifts S to a velocity z and rebins the result to the dispersion of G . Thus $\mathcal{L}(T)$ is now the filtered version of the redshifted template. Even with modest contemporary computing resources, one does not require that G be binned logarithmically in wavelength, as template spectra with “infinite” signal-to-noise ratios can be quickly resampled to match the wavelength sampling of the observed galaxy spectrum.

The above representations explicitly weight each pixel uniformly, and we now expand the cross-correlation to include non-uniform weighting:

$$\mathcal{C}(z) = \sum_{i=1}^n \frac{M_i}{\epsilon_i^2} \mathcal{L}(T)_i \mathcal{L}(G)_i \quad (3)$$

where ϵ is the noise in $\mathcal{L}(G)$. M is a mask spectrum, equal to unity for pixels to be included in the filtering and computation of the cross-correlation, and equal to zero for those pixels to be excluded. The use of M allows for the explicit masking of emission lines, uncorrected (or poorly corrected) telluric absorption, and bad pixels.

In previous incarnations of the cross-correlation method the filtering function, \mathcal{L} , eliminates low frequencies from the power spectrum (Tonry & Davis 1979; Kurtz & Mink 1998; Glazebrook, Offer & Deeley 1998). Such FFT-based procedures, like Eq. 2, weight each pixel uniformly, and are unsuitable when there are strongly discrepant features or highly localized sources of noise. Instead we choose to fit a low-order, non-parametric function, weighting each pixel by M_i/σ_i^2 , where σ^2 is the variance in G . Weighting each pixel in this way is the source of the real power of this technique over methods that operate with Fourier transforms of the data.

We choose to define \mathcal{L} as an operation that fits cubic B-splines (de Boor 1978; Dierckx 1993) to T and G , using pixel weights of M/σ^2 in performing the fits to both spectra. B-splines were chosen because of their non-parametric

nature, but one could also adopt polynomials or sines and cosines. While sines and cosines can be used to better approximate Fourier filtering, we prefer B-splines because they are more efficiently computed. In fitting the B-splines, knots are selected to be widely spaced, with typical *restframe* spacings of 33\AA^2 . The knot spacing should be large enough to allow the fit to follow the continuum shape, without fitting the detailed features of a given spectrum. Note that $\mathcal{L}(G)$ and $\mathcal{L}(T)$ must be recomputed at every radial velocity one wishes to sample with $\mathcal{C}^w(z)$. This is because as one scans through different radial velocities M/σ^2 weights different locations in the template’s intrinsic wavelength range. For example, as one re-samples the template over a broad range of redshifts, an excess of localized noise, e.g., from subtraction of the 5577\AA night-sky emission line, may bias the continuum fit in the galaxy, and this bias must be reproduced in the continuum fit to the template. By redoing the fit at every velocity, the template and galaxy are thus filtered in exactly the same way, with each pixel contributing the same weight in the fit to the galaxy spectrum and in the fit to the template spectrum. Because the fitting of B-splines is computationally inexpensive, the extra burden is not significant.

The B-spline fits to T and G are represented by T' and G' . Thus $\mathcal{L}(G) \equiv G/G' - 1$ and $\mathcal{L}(T) \equiv T/T' - 1$ such that $\mathcal{L}(G)$ and $\mathcal{L}(T)$ have zero mean. Note too that ϵ^2 , the variance in $\mathcal{L}(G)$, is well approximated by $\epsilon^2 \approx \sigma^2 G^{-2}$, as long as $\kappa^{1/2} \gg 1$, where κ is equal to the knot spacing (in pixels). Because our B-spline knots have fixed spacing in restframe wavelengths, and because masked galaxy pixels are coincident with different redshifted template pixels, T' and G' themselves are mildly sensitive to v . However, because M/σ_i^2 was used to weight the B-spline fits to both T and G , the sensitivities of T' and G' to z should be similar, if not identical.

The inverse-variance weighted cross-correlation, \mathcal{C}^w , can now be written explicitly as

$$\mathcal{C}^w(z) = \sum_{i=1}^n \frac{M_i}{(\sigma_i/G'_i)^2} \left(\frac{G_i}{G'_i} - 1 \right) \left(\frac{T_i}{T'_i} - 1 \right) \quad (4)$$

$$= \sum_{i=1}^n \frac{M_i}{\sigma_i^2} (G_i - G'_i) (T_i^\circ - G'_i) \quad (5)$$

where $T^\circ = T(G'/T')$ is a “continuum-matched” version of the template spectrum, similar to that found in techniques for measuring velocity dispersions (Rix & White 1992; Kelson et al. 2000). This expression for the inverse-variance weighted cross-correlation (Eq. 5) illustrates the key advantage of this method: the down-weighting of noisy data and the elimination of specific features, either due to detector defects, non-Poisson sources of noise, or unwanted features intrinsic to the galaxy and template spectra.

To demonstrate an important consequence of Eq. 5, we recall the minimization of χ^2 in the direct fitting method for the measurement of velocity dispersions (Kelson et al. 2000):

$$\chi^2 = \sum_{i=1}^n \frac{M_i}{\sigma_i^2} \{G_i - I_i[B \circ S]_i - P_i\}^2 \quad (6)$$

² In order to reproduce the fit of a single sine wave, three knots are required. Therefore, the equivalent bandpass in the Fourier domain is a factor of three larger, or $\sim 100\text{\AA}$. This bandpass is appropriate for galaxy spectra dominated by stellar populations. For quasars, we employ a large enough bandpass to encompass the broad emission lines.

where B , the line-of-sight velocity distribution, is a function of velocity and velocity dispersion, S is a template spectrum, I is a low-order multiplicative function that includes both the mean line strength differences between G and S , as well as differences between the instrumental responses in G and S , and P is a low-order additive continuum function. Since we are only interested in measuring radial velocities, we can simplify the expression with the following definition:

$$S^\circ = I(B \circ S) + P \quad (7)$$

Thus

$$\chi^2 = \sum_{i=1}^n \frac{M_i}{\sigma_i^2} \{G_i - S_i^\circ\}^2 \quad (8)$$

$$= \sum_{i=1}^n \frac{M_i}{\sigma_i^2} [G_i^2 + S_i^{\circ 2} - 2G_i S_i^\circ] \quad (9)$$

Expanding Eq. 5

$$\mathcal{C}^w(z) = \sum_{i=1}^n \frac{M_i}{\sigma_i^2} [G_i T_i^\circ + G_i'^2 - G_i G_i' - T_i^\circ G_i'] \quad (10)$$

The summations over the last three terms of Eq. 10 and the first two terms of Eq. 9 have only a second-order dependence on z near the peak of the cross-correlation. Thus,

$$\frac{d\mathcal{C}^w(z)}{dz} = \frac{d}{dz} \sum_{i=1}^n \frac{M_i}{\sigma_i^2} G_i T_i^\circ \equiv -\frac{1}{2} \frac{d\chi^2}{dz} \quad (11)$$

Therefore the shape and height of the peak of the inverse-variance weighted cross-correlation can be straightforwardly employed to infer confidence limits on the measured radial velocities: $-\Delta\mathcal{C}^w$ behaves identically to $\Delta\chi^2$ with a single degree of freedom (the radial velocity). Put more simply, the determination of the $\pm 1\text{-}\sigma$ confidence limits only requires solving for $z_{\pm 1\sigma}$ such that

$$\mathcal{C}^w(z_{\pm 1\sigma}) \equiv \mathcal{C}^w(z_{peak}) - \frac{1}{2} \quad (12)$$

Equations 11 and 12 thus provide the second key advantage to measuring redshifts using the inverse-variance weighted cross-correlation: the formal errors in the redshift are trivial to derive (compare to Tonry & Davis 1979; Heavens 1993).

Note that the value at the peak of the weighted cross-correlation, $\mathcal{C}^w(z_{peak})$, is proportional to the square of the signal-to-noise ratio. This fact aids ones intuition into the behavior of Eq. 12. For data with low signal-to-noise ratios, the numerical search for the $\pm 1\sigma$ radial velocities must clearly travel a greater distance from z_{peak} . While accurate recovery of the errors requires knowledge of the true variance, under- or over-estimation of the variance can be accounted for by rescaling the errors by an ‘‘effective’’ reduced- χ^2 , which we refer to as $(\chi_{eff}^2)^{1/2}$, where

$$\chi_{eff}^2 = \sum_{i=1}^n \frac{M_i}{\sigma_i^2} (G_i - T_i^\circ)^2 / \left[\left(\sum_{i=1}^n M_i \right) - \kappa - 1 \right] \quad (13)$$

In a χ^2 -sense, there are $\nu = (\sum_{i=1}^n M_i) - \kappa - 1$ degrees of freedom, where $\sum_{i=1}^n M_i$ is the number of pixels used in the computation of the cross-correlation, and there are $\kappa + 1$ parameters constrained by the data (the knot coefficients and the redshift).

For data with low-to-moderate signal-to-noise ratios χ_{eff}^2 is dominated by Poisson statistics, and thus the formal errors from Eq. 12 are correctly rescaled to account for incorrectly estimating the noise spectrum. For data with higher signal-to-noise ratios, χ_{eff}^2 is dominated by mismatch between the galaxy and template spectra. In order to address the effects of template mismatch on our measurements of z , and on our estimates of the redshift errors, we now proceed to a discussion of Monte Carlo simulations in which we explicitly test the validity of the measurements and their errors over a range of signal-to-noise ratios.

3. SIMULATIONS

In the previous section, we showed that the weighted cross-correlation is simple to compute, and that the formal errors in the location of the peak are also simple to derive. In this section we generated 140 noisy realizations of a model spectrum of a 6 Gyr old single burst stellar population (Bruzual & Charlot 2003) at $z = 0.3$, creating 20 spectra at each of the signal-to-noise ratios $S/N \in \{2, 4, 8, 12, 16, 20, 24, 28, 32, 48, 64, 96, 128\}$ (per 5Å pixel). Prior to adding noise, each spectrum was convolved with a Gaussian of dispersion 250 km/s and the resolution was degraded to that of our Abell 2550 data (see §4). Poisson noise was added to each spectrum equal to the noise expected from both the object and the sky as well as an electronic read noise of $5e^-$. The sky spectrum was taken from our observations of Abell 2550. Using these artificially noisy spectra, we used 6 Gyr-old, 2 Gyr-old, and 1 Gyr-old model spectra as templates, utilizing the weighted cross-correlation method of §2 to measure redshifts for the noisy simulated data. Note that the template spectra that we employ have been convolved to a nominal resolution of 12Å (FWHM), though with the 1Å standard Bruzual & Charlot (2003) sampling. This resolution approximates the simulated galaxy data to zeroth order.

Figure 1 shows the accuracy of the results one obtains when the 6 Gyr-old template is used to measure the redshifts of the simulated galaxy spectra. In Fig. 1(a) we show the the formal errors plotted against the peak value of the inverse-variance weighted cross-correlation, $\mathcal{C}^w(z_{peak})$. Because $\mathcal{C}^w(z_{peak})$ is proportional to the square of the mean signal-to-noise ratio, the formal error in the radial velocity should follow $\sigma_z \propto \mathcal{C}^w(z_{peak})^{-1/2}$. The dashed line shows a line with that expected slope, shifted vertically to pass through the median of the results of the simulations. A least-squares fit confirms our expectation at the 2% level. Fig. 1(b) shows that χ_{eff}^2 remains nearly constant (\sim unity), with a small increase with signal-to-noise ratio. This increase occurs because of the difference between the resolutions of the template and the simulated galaxy spectra. In Fig. 1(c-e) we plot histograms of the measurement errors, normalized by the formal errors as reported by Eq. 12 and rescaled by Eq. 13. In these three panels the results have been binned by the signal-to-noise ratios of the simulated galaxy spectra, in order to demonstrate the accuracy of the reported errors over a broad range of ‘‘data’’ quality. When the correct template has been used the measurement errors are distributed in a manner consistent with a Gaussian whose centroid is identical to zero and whose standard deviation is equal to

the formal errors (to within the statistics of the simulations). This result is independent of the signal-to-noise ratios of the spectra, when signal-to-noise ratios are sufficiently high that the cross-correlation has a unique peak. In our simulations, using the 33Å knot-spacings, the cross-correlations for nearly all of the spectra with $S/N = 2$ per pixel had multiple, equally statistically significant peaks. About 1/3 of the spectra with $S/N = 4$ per pixel suffered from similar uniqueness problems. When $S/N > 8$ per pixel, the cross-correlations were no longer plagued with multiple, significant peaks. However, when we relaxed the knot-spacing to 100Å or greater, the cross-correlations at lower signal-to-noise ratios became more robust. In general when one must analyze data of poor quality, one should consider using significantly larger knot spacings.

In Figures 2 and 3, we plot the results obtained with 2 Gyr-old and 1 Gyr-old template spectra, respectively. Fig. 2(a) and 3(a) show that the scaling of the formal error with $\mathcal{C}^w(z_{peak})$ breaks down at high signal-to-noise ratios, when the mismatch between the template spectra and the simulated galaxy spectra dominates over the Poisson noise. This point is shown more clearly by the correlation of χ_{eff}^2 with $\mathcal{C}^w(z_{peak})$ in Figs. 2(b) and 3(b).

At higher signal-to-noise ratios, systematic errors due to template mismatch begin to dominate over the measurement errors due to Poisson statistics. Template mismatch becomes important when $(\chi_{eff}^2)^{1/2} \gg 1$ and the mean reported error is approximately equal to the mean systematic error. Figures 2(c-e) and 3(c-e) show that for $S/N > 15$, the measurement errors diverge from the expected distribution. The centroid reflects the systematic error but its magnitude is equivalent to the error reported by Eqs. 12 and 13. The resulting inflation of the errors is also shown by the narrowing of the histograms of normalized measurement errors. The standard deviations are significantly less than unity. In the regime of higher signal-to-noise ratios, when the systematic effects of template mismatch can be important, the errors, inflated by $(\chi_{eff}^2)^{1/2}$, are larger than one would expect from random noise alone. Note that the direction of the systematic error will depend on the nature of the mismatch between the galaxy template, and our tests only used templates that were of the same age or younger than the simulated galaxies. Had we adopted a young SED for the simulated galaxy spectra, the systematic errors shown in Figures 2(c-e) and 3(c-e) would have the opposite signs.

As many authors have noted, the mismatch of the template with the underlying spectrum of the observed galaxy can seriously affect the results from published cross-correlation methods (e.g. Tonry & Davis 1979; Kurtz & Mink 1998) and, as mentioned in §1, Glazebrook, Offer & Deeley (1998) attempted to reduce the effects of template mismatch by employing multiple template spectra simultaneously. Using our method, one may simply reduce template mismatch errors by selecting the template that (a) minimizes σ_z , or (b) provides a reasonable match between $\mathcal{L}(T)$ and $\mathcal{L}(G)$. For lower quality spectra, errors due to template mismatch are not statistically significant.

Based on our simulations, we conclude that the solution to Eq. 12 accurately recovers the measurement errors. We now proceed to the first application of our method to real data obtained in our survey of Abell 2550.

4. APPLICATION TO ABELL 2550

As part of our on-going survey for X-ray detected AGN in clusters (Martini, Kelson, Mulchaey, & Trager 2002), the field of Abell 2550 was targeted with LDSS2 on the Baade 6.5m telescope at Magellan. Abell 2550 is a Richness Class 2 cluster (Abell 1958) with a published redshift of $z = 0.12260$ (Caretta et al. 2002). We observed this field with three slit-masks, with total exposure times of 2700 s, 6000 s, and 6000 s. The grism and slit-width provided a resolution of $R \approx 400$. The slit-masks were designed to include optical counterparts to unresolved *Chandra* sources to $R_c \leq 20.5$ mag, and an additional sample of optically detected galaxies to $R_c \leq 21$ mag. A total of 40 objects were observed; 8 of these contained unresolved *Chandra* point sources within 1''5 of the optical centers of the galaxies.

4.1. The Data

The data were preprocessed using standard procedures. The treatments of the y -distortion, the rectifications, and the subtraction of the two-dimensional sky background were performed using algorithms discussed in Kelson et al. (2000) and Kelson (2003). The extractions were performed using a new method whereby one fits a given object's pixels, from every exposure, with a single bivariate B-spline. This B-spline becomes the internal representation of the object's spectrum as a function of wavelength and spatial position. By using the pixels from multiple exposures, a single B-spline representation of each object can be constructed in a way that fully utilizes the sub-pixel sampling provided by shifts in the instrument that occur between integrations. In this way, cosmic-rays and other discrepant pixels can also be straightforwardly rejected in the extraction procedure. Once the fit of the two-dimensional B-spline has been performed to a given object, a one-dimensional spectrum can be computed at any set of wavelengths desired for the extracted spectrum. This computation is inexpensive because integration of the bivariate B-spline over the desired aperture is a simple procedure (Dierckx 1993). Further discussion of this method is beyond the scope of this paper and will be presented in greater detail at a later date.

The 36 extracted spectra for the galaxies for which we could measure redshifts are shown in Figure 4. Of the 40 original targets, one was a foreground M star and we do not include it in these discussions. Three additional sources had insufficient signal-to-noise to provide unique and statistically meaningful redshift determinations. Because the masks were exposed multiple times, we measured redshifts from the individual exposures, tested our estimates of the measurement errors, and found them to be accurate.

Figures 5 through 9 show detailed examples from our weighted cross-correlation analysis. The top panels show observed galaxy spectra, G (black), the B-spline fit to the galaxy continuum, G' (blue), the B-spline fit to the redshifted, rebinned template's continuum, T' (green), and the "continuum-matched" template spectrum, T° (red). The templates are shown redshifted to z , the peak redshift of the weighted cross-correlations. In the middle panels, the continuum-filtered spectra $\mathcal{L}(G)$ (black) and $\mathcal{L}(T)$ (red). The error in $\mathcal{L}(G)$, approximated as σ/G' , is

shown in blue. The A- and B-band atmospheric absorption features are clearly visible because they were masked in the B-spline filtering. These features were masked by M in the summation over pixels to compute $C^w(z)$, shown in the bottom panels. In the bottom panels, the coarsely sampled weighted cross-correlations are shown and the insets show the regions of the peak of C^w in high-resolution, with horizontal lines indicating the ± 1 -, ± 2 -, and ± 4 - σ confidence limits. Note that for spectra with lower signal-to-noise ratios the peaks are broader, leading to larger estimates of the formal errors. In these examples, we also specifically masked the region around the 5577Å night-sky emission line, and for a few galaxies we also masked the region around the 6300Å night-sky emission line. While the cross-correlation procedure correctly down-weights the data at the locations of these features, we masked these features for cosmetic reasons.

4.2. The Redshifts

Table 1 lists our galaxy identifications along with their right ascensions and declinations. Along with the measured redshifts we also list whether or not the redshift was measured using the cross correlation technique alone, or was measured using Gaussian fits to emission lines after a cross correlation had been performed (using either a star-forming galaxy spectrum or the Sloan Digital Sky Survey composite quasar spectrum as templates in the cross-correlation; Kennicutt 1992; Vanden Berk et al. 2001). Note that the quoted redshifts and errors are those derived by the techniques discussed in this paper and they have not been transformed to the frame of the cosmic microwave background. In Figure 4, we label many of the features seen in the individual spectra.

Figure 10 shows a histogram of the 36 redshifts. While the coarsely-binned histogram indicates the presence of a large structure at the “known” redshift of the cluster, $z = 0.12$ (Caretta et al. 2002), the high-resolution inset shows that there are two galaxy groups near that redshift. The brightest galaxy in the group, #422, is at the redshift given by Caretta et al. (2002) and we also note that the extended X-ray emission in the archived Chandra imaging is centered there. Out of the 36 galaxies for which we have redshifts, there is one group of 5 galaxies centered at $z = 0.1215 \pm 0.0003$ with a velocity dispersion of $\sigma = 290 \pm 290$ km/s. (using the bi-weight estimator for location and scale; Beers, Flynn, & Gebhardt 1990). The second low-redshift group contains 7 galaxies and is located at $z = 0.1090 \pm 0.0009$ with a velocity dispersion of $\sigma = 600 \pm 250$ km/s. The > 3000 km/s velocity separation between these two groups indicates that they are not physically associated with each other.

Nevertheless, neither structure dominates the galaxy populations in the field of the Chandra imaging of the cluster and it appears as though there are actually 4 low redshift groups along the line of sight, extending to $z = 0.2$. Aside from the two groups described above, a third group of 5 galaxies appears at $z = 0.158 \pm 0.003$ with $\sigma = 950 \pm 915$ km/s, and fourth group of 7 galaxies appears at $z = 0.2014 \pm 0.0003$ with $\sigma = 210 \pm 150$ km/s (the errors in the redshifts are standard errors of the mean). Based on our sample, we conclude that Abell 2550 itself is misclassified as Richness Class 2. The apparent richness of

the cluster is most likely due to a superposition of four groups below $z \lesssim 0.2$ along the line-of-sight. The specifics of the X-ray detected galaxies will be discussed in a future paper.

5. SUMMARY

We have presented a new technique for measuring redshifts using the cross-correlation in the Real domain. We have derived this method because Fourier-based cross-correlations treat each datum uniformly even though observations have non-uniform sources of noise and other systematic defects. These common problems are not easily dealt with when one measures redshifts using Fourier-based techniques because such methods explicitly weight the pixels in a given spectrum uniformly (e.g. Tonry & Davis 1979; Kurtz & Mink 1998). Fortunately performing cross-correlations in the Fourier domain is no longer computationally necessary, and one is now free to introduce non-uniform weighting to the calculations. We opt to weight by the inverse of the variance. There are two important reasons for doing so: first, one can explicitly ignore pixels that are heavily contaminated, and second, because those pixels included in the computation have been weighted by the inverse of the variance, the units of the weighted cross-correlation are exactly half that of χ^2 . Thus, the topology of the peak of the weighted cross-correlation can be straightforwardly interpreted to estimate confidence limits on the measured redshifts.

We tested the accuracy and precision of this method using simulated spectroscopic observations. We verified that the formal errors accurately reflect the true measurement errors, when the correct template is used. For low signal-to-noise data, errors due to template mismatch are negligible. For spectra with high signal-to-noise ratios (in our artificial spectra with dispersions of 5\AA per pixel, this constitutes $S/N \gtrsim 36$ per pixel), template mismatch is statistically significant. However, one can minimize it by selecting the template that provided the smallest formal error (or the lowest χ_{eff}^2). Nevertheless, we found that while the contribution to the error due to photon statistics continued to shrink with increasing signal-to-noise ratio, the systematic error due to mismatch was well approximated by the reported errors, when rescaled by $(\chi_{eff}^2)^{1/2}$.

Lastly, we applied the method to a portion of our survey data, specifically data in the field of the cluster Abell 2550. We obtained 36 redshifts; five in a group centered at the published redshift of the cluster, $z = 0.12$, and another seven in a slightly lower redshift group at $z = 0.109$. In total, we found four groups of galaxies between redshifts of $z = 0.1$ and $z = 0.2$. We conclude that the “cluster” only appears rich because the line-of-sight contains a superposition of four populated groups spanning redshifts from $z = 0.10$ to $z = 0.20$. However, Caretta et al. (2002) suggest that Abell 2550 is part of a large super-cluster in Aquarius and two of these four groups we detected may indeed be members of that superstructure.

In conclusion, we have successfully derived a robust, Real-domain variant of the cross-correlation method for the measurement of redshifts. The method is easy to implement, with the added benefit that errors in the measured redshifts are also easy to derive with no particularly cumbersome machinery required. While we have discussed

the weighted cross-correlation in the context of extragalactic redshift surveys, there are no known barriers to applying the method to the measurements of the radial velocity of stars or other astronomically interesting targets.

The first author would like to thank the staff at the Carnegie Observatories for their continued support. Furthermore, we acknowledge Stephen Helsdon, guinea pig extraordinaire, for extensively testing our implementation

of the algorithm presented in this paper. Support for this work was provided by the National Aeronautics and Space Administration through Chandra Award Number 4700793 issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-390. PM was supported by a Carnegie Starr Fellowship.

REFERENCES

- Abell, G. O. 1958, *ApJS*, 3, 211
 Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, 100, 32
 Bruzual, G. & Charlot, S. 2003, *MNRAS*, in press
 Caretta, C. A., Maia, M. A. G., Kawasaki, W., & Willmer, C. N. A. 2002, *AJ*, 123, 1200
 Cooley, J. W. and Tukey, J. W., 1965, *Mathematics of Computation*, 19, 297
 de Boor, C., *Applied Mathematical Sciences*, New York: Springer, 1978
 Dierckx, P., *Curve and Surface Fitting with Splines*, New York: Oxford University Press, 1993
 Glazebrook, K., Offer, A. R., & Deeley, K. 1998, *ApJ*, 492, 98
 Griffin, R. F. 1967, *ApJ*, 148, 465
 Heavens, A. F. 1993, *MNRAS*, 263, 735
 Kelson, D. D. 2003, *PASP*, 115, 688
 Kelson, D. D., Illingworth, G. D., van Dokkum, P. G., & Franx, M. 2000, *ApJ*, 531, 159
 Kennicutt, R. C. 1992, *ApJS*, 79, 255
 Kurtz, M. J. & Mink, D. J. 1998, *PASP*, 110, 934
 Martini, P., Kelson, D. D., Mulchaey, J. S., & Trager, S. C. 2002, *ApJ*, 576, L109
 Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. *Numerical Recipes*, New York: Cambridge University Press, 1986
 Rix, H.-W. & White, S.D.M. 1992, *MNRAS*, 254, 389
 Simkin, S. M. 1974, *A&A*, 31, 129
 Tonry, J. & Davis, M. 1979, *AJ*, 84, 1511
 Vanden Berk, D. E. et al. 2001, *AJ*, 122, 549

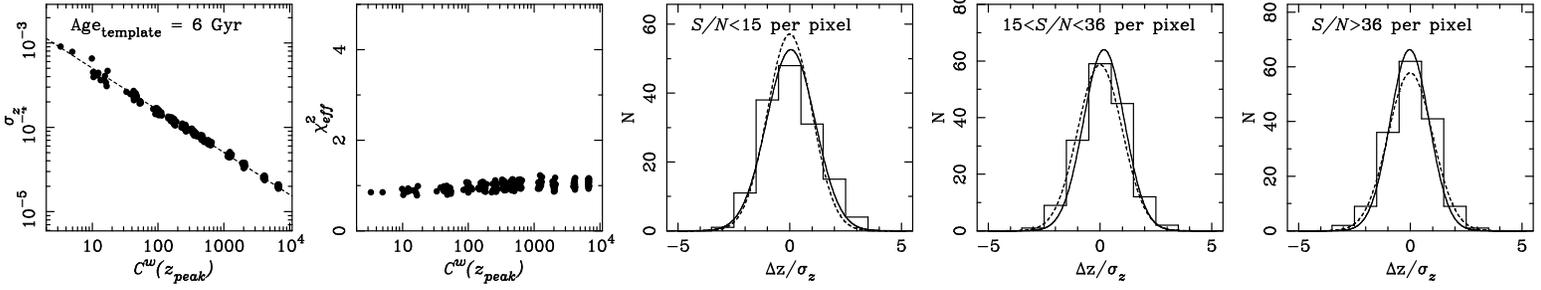


FIG. 1.— Redshift measurement errors as a function of signal-to-noise ratio for the correct template (6 Gyr SSP). (a) Redshift uncertainty vs. height of the cross-correlation peak $C^w(z_{peak})$. The errors are expected to follow $\sigma_z \propto C^w(z_{peak})^{-1/2}$, shown by the dashed line (normalized to match the results of the simulation). (b) The correlation of χ_{eff}^2 with $C^w(z_{peak})$. (c) The histogram of measurement errors, scaled by the estimates of formal errors, for those spectra with signal-to-noise ratios $S/N < 15$ per pixel. The dashed line shows the expected Gaussian with standard deviation of unity and the solid line shows the Gaussian that best represents the “observed” distribution of errors; (d) same as (c) for $15 < S/N < 36$ per pixel; (e) same as (c) for $S/N > 36$ per pixel.

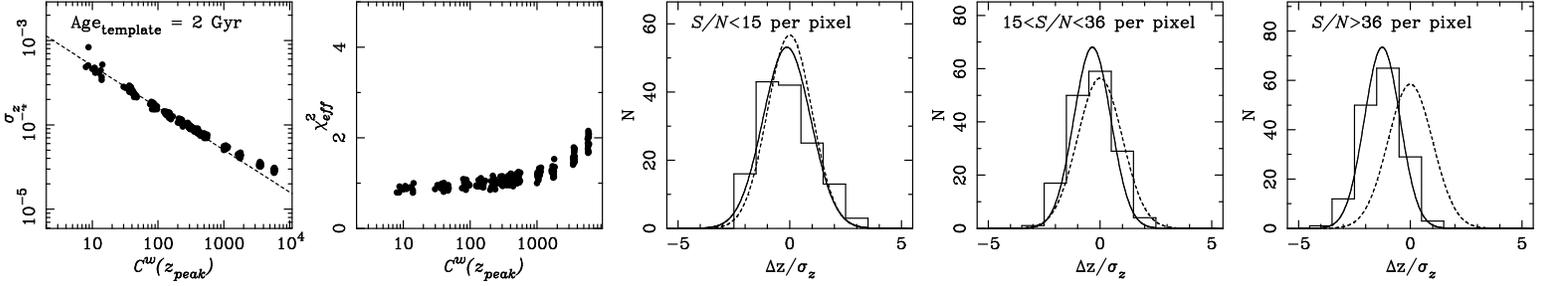


FIG. 2.— Same as Figure 1 for a mismatched template (2 Gyr SSP). (a) the redshift uncertainties no longer follow $C^w(z_{peak})^{-1/2}$. At low signal-to-noise ratios the uncertainties are still dominated by Poisson statistics and $\sigma_z \propto C^w(z_{peak})^{-1/2}$, as in Figure 1(a); however, at high signal-to-noise ratios, the systematic errors due to template mismatch inflate χ_{eff}^2 , shown in (b), causing σ_z to diverge from the expected relation. Note the shift in the centroids of the distributions in (c) through (e) as template mismatch becomes the dominant source of error. One can minimize the effects of template mismatch, at any signal-to-noise ratio, when one minimizes χ_{eff}^2 , and thus minimizes σ_z .

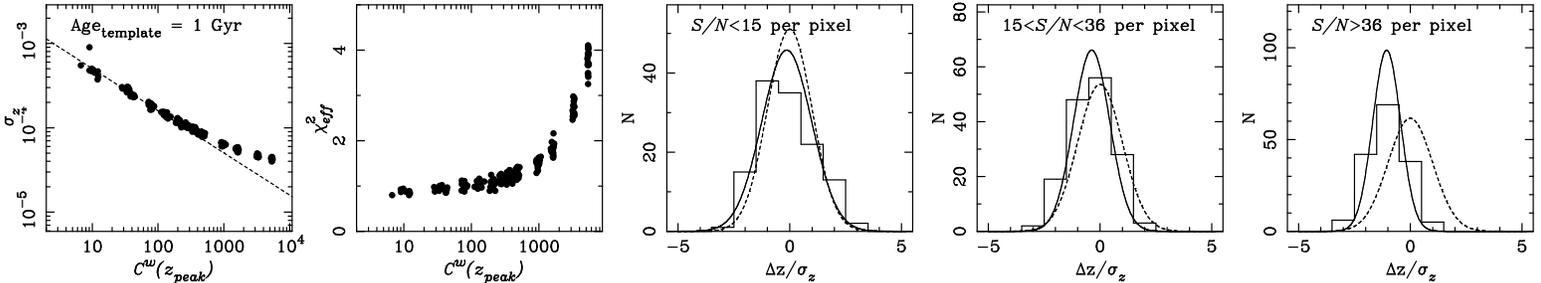


FIG. 3.— Same as Figure 2 but for a 1 Gyr SSP template.

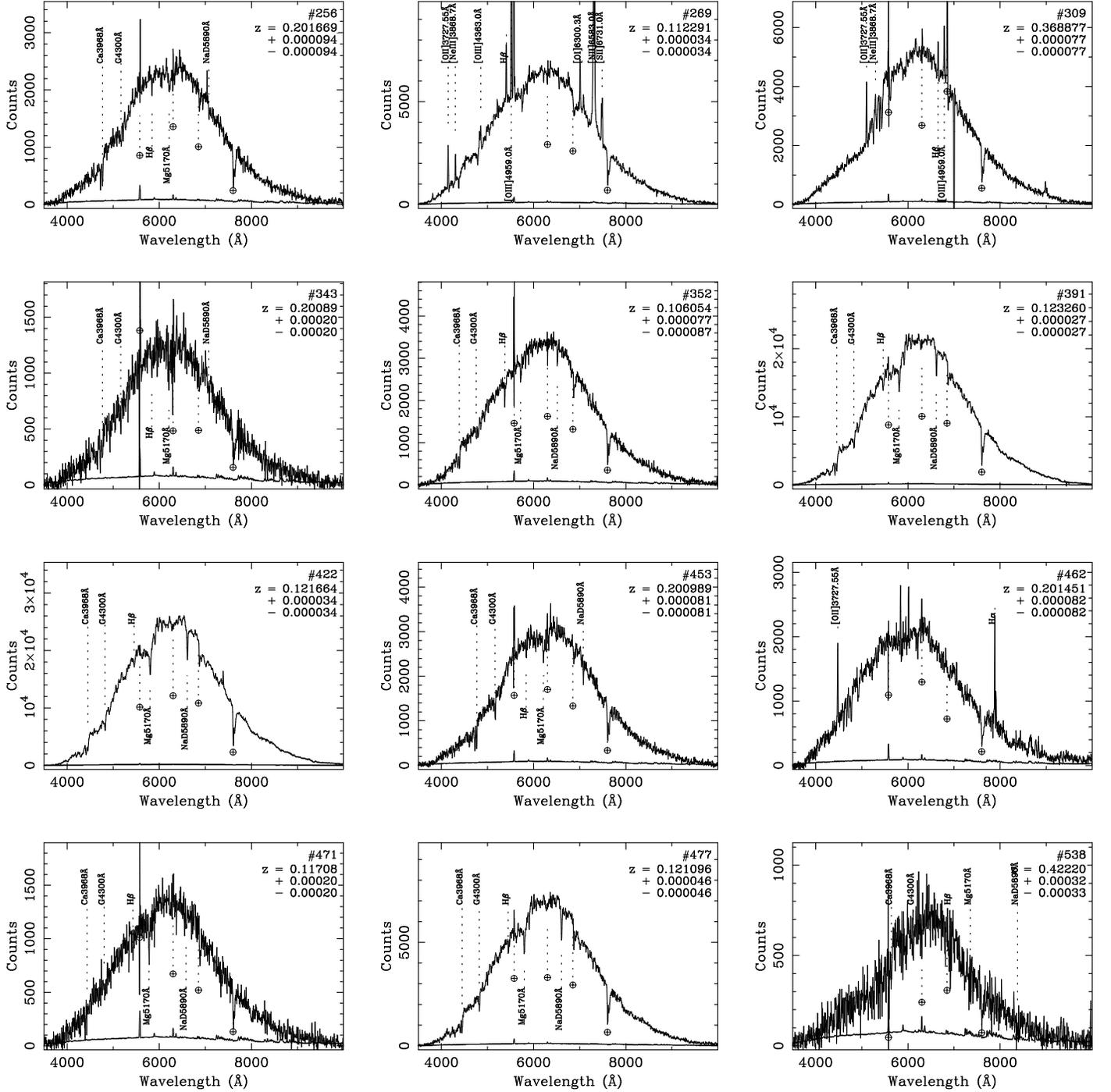


FIG. 4.— The one-dimensional extracted spectra for the 36 galaxies for which we were able to obtain redshifts. These spectra have not been flux-calibrated; this has no effect on the results. Some of the more common spectral features are also labeled at their observed wavelengths.

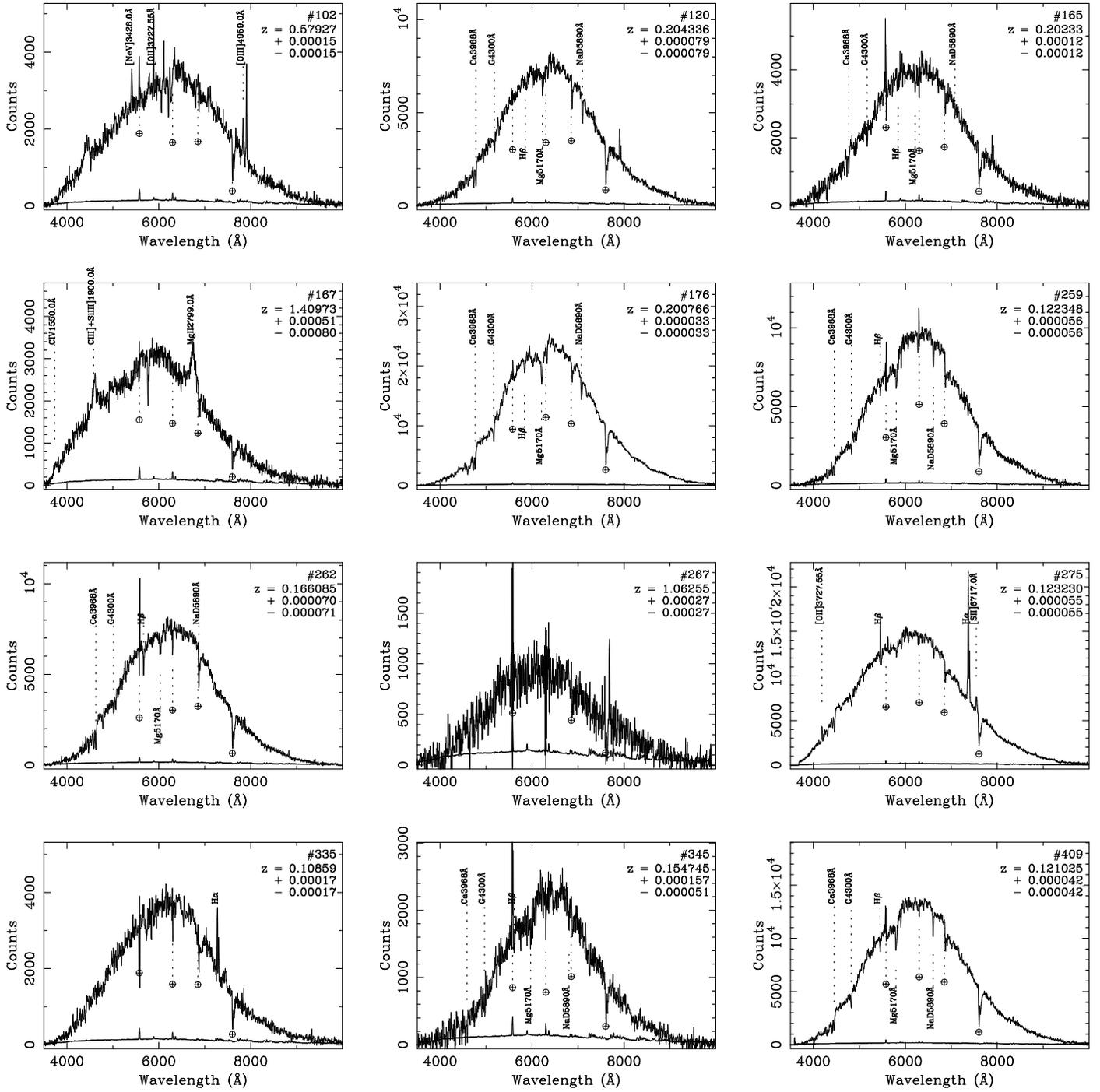


Fig. 4 continued.

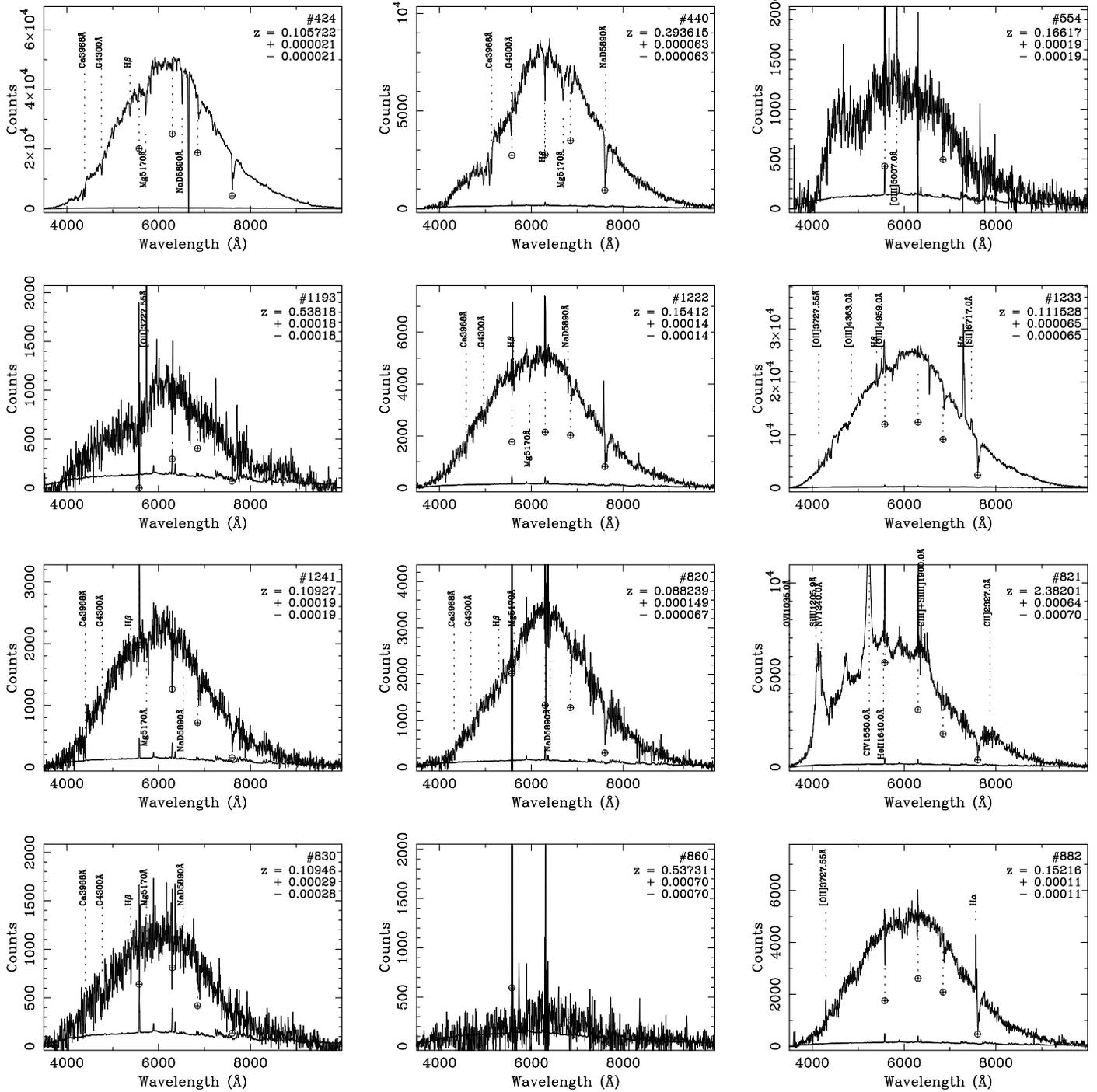


Fig. 4 continued.

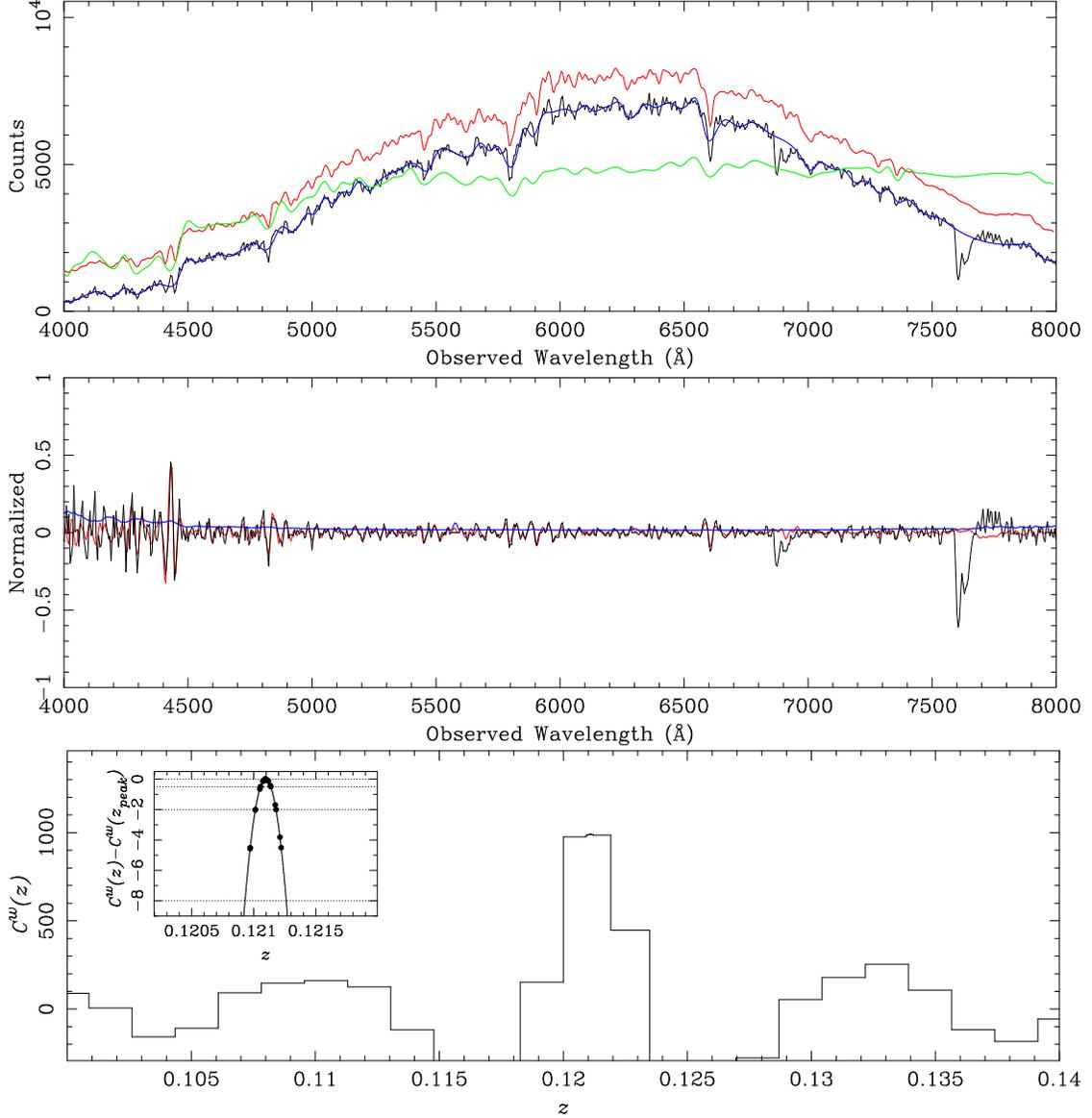


FIG. 5.— Weighted cross-correlation results for galaxy #477. (top) The observed spectrum, G , is shown in black. The low-order B-spline fit, G' , is shown in blue. The low-order B-spline fit to the template spectrum, T' , is shown in green. The “continuum-matched” version of the template, T° is shown in red. The case shown is for the redshift at the peak of the inverse-variance weighted cross-correlation (see bottom panel). Note that regions affected by the atmospheric A- and B-band absorption have been given zero weight in the fitting and computation of C^w . (middle) The continuum-filtered spectrum of the galaxy, $\mathcal{L}(G)$, and the filtered template, $\mathcal{L}(T)$, are shown in black and red, respectively. The blue line shows the expected noise in $\mathcal{L}(G)$, σ/G' . (bottom) The coarsely sampled inverse-variance weighted cross-correlation is shown for a range of redshifts $0.1 \leq z \leq 0.14$. The inset figure shows the tip of the peak at high-resolution. The horizontal lines indicate offsets from the peak $\Delta C^w \in \{1/2, 2, 8\}$, corresponding to the confidence limits of ± 1 - ± 2 - and $\pm 4\sigma$, respectively.

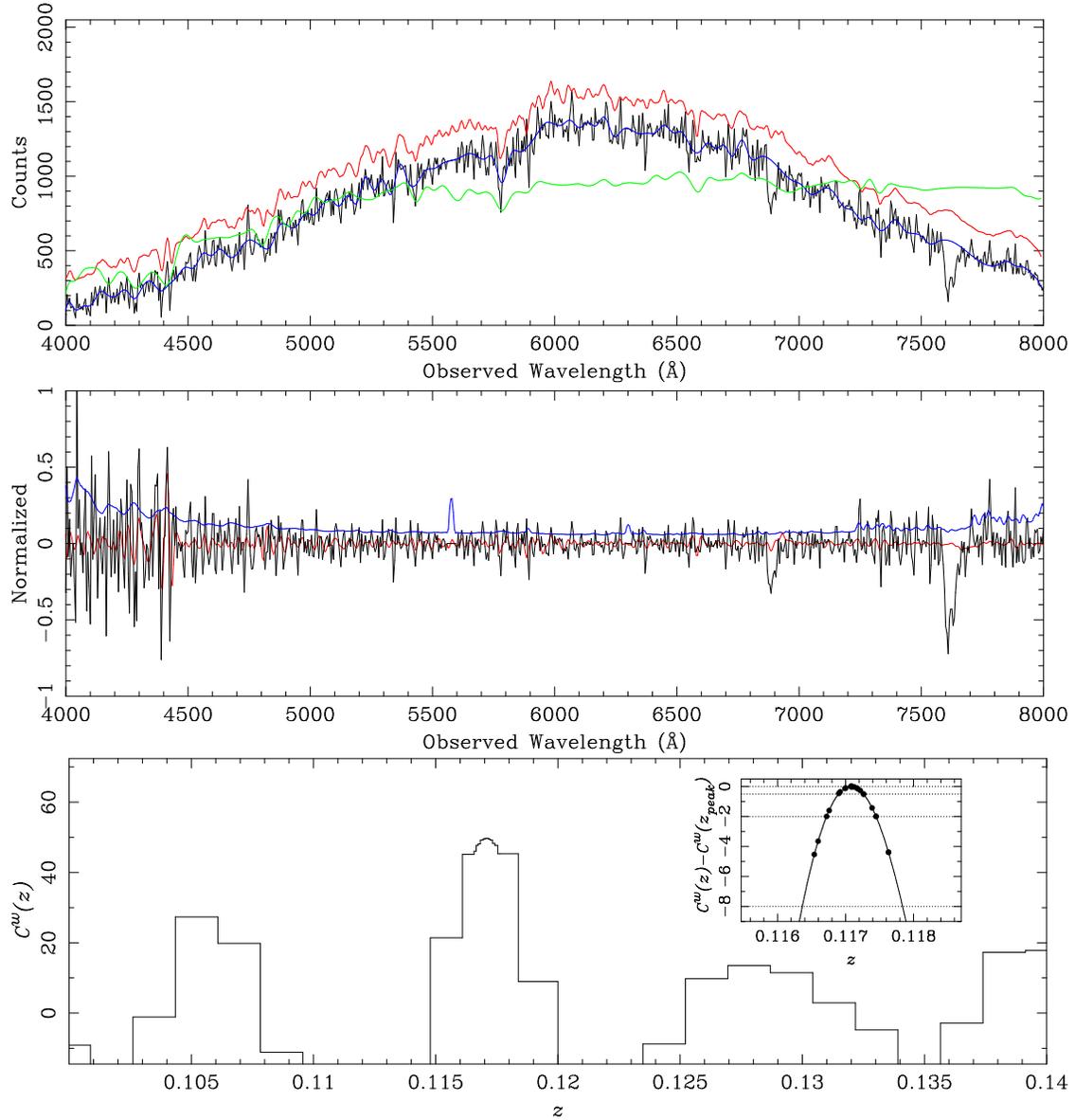


FIG. 6.— Same as Figure 5 for galaxy #471.

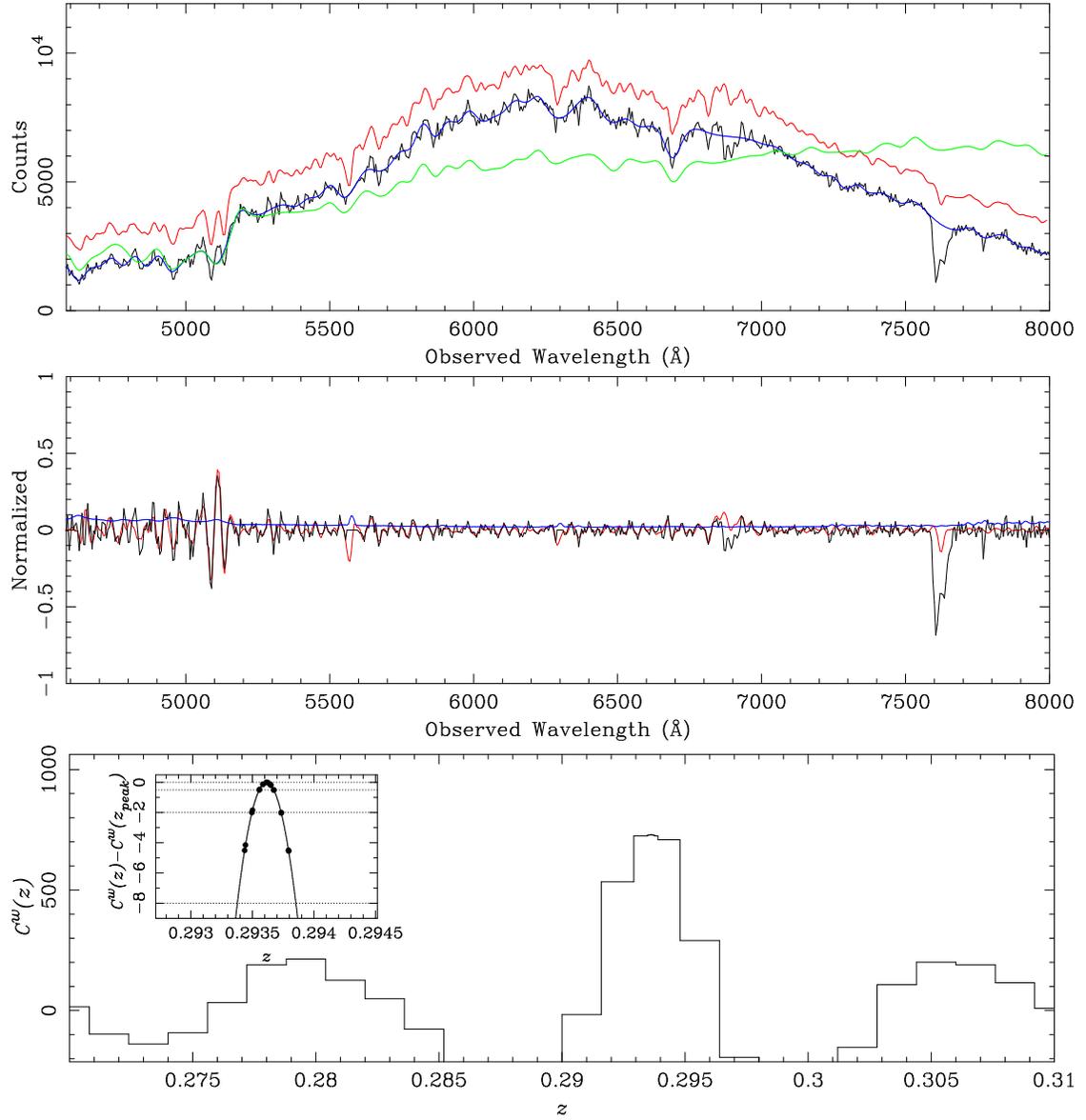


FIG. 7.— Same as Figure 5 for galaxy #440.

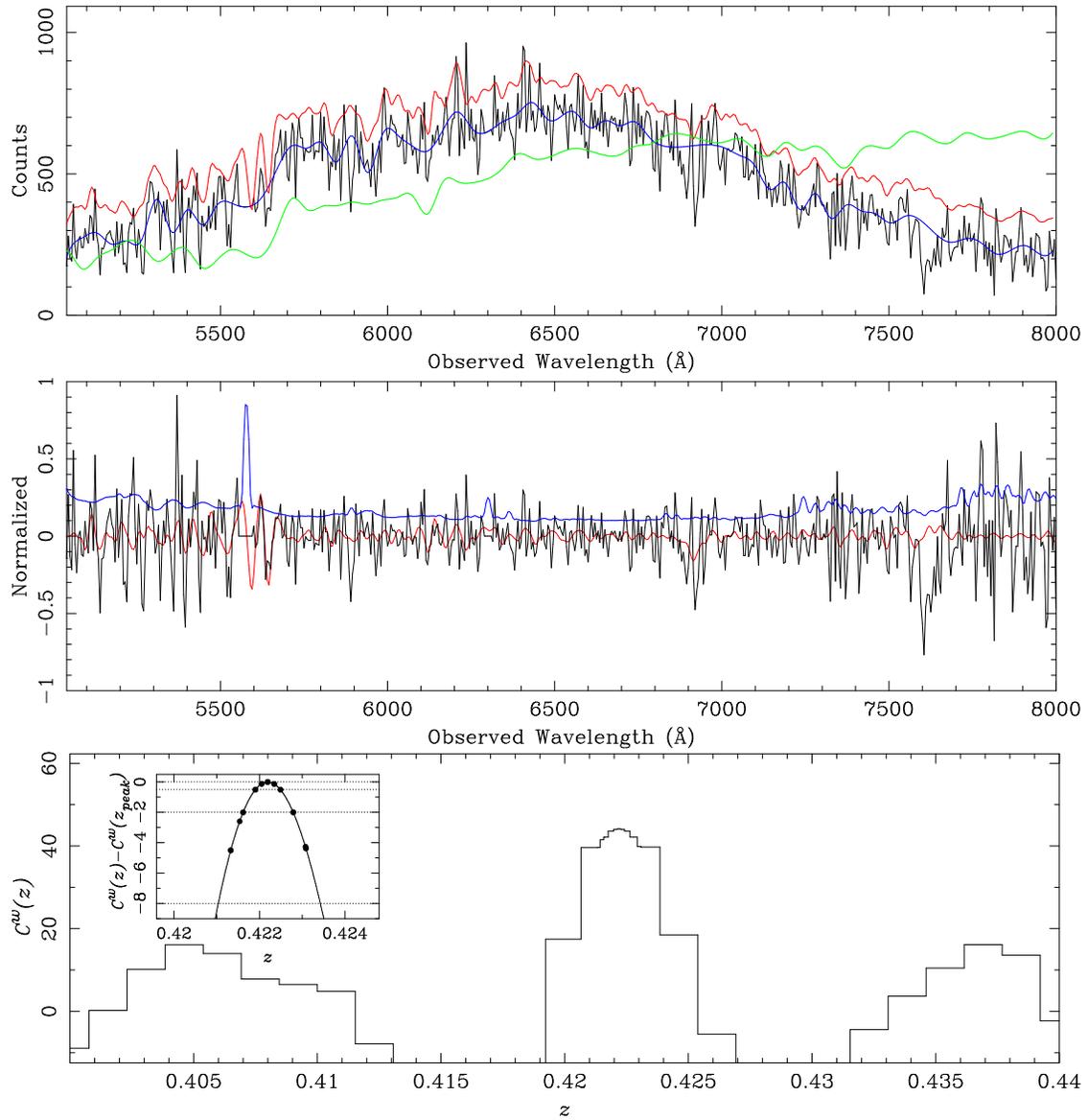


FIG. 8.— Same as Figure 5 for galaxy #538.

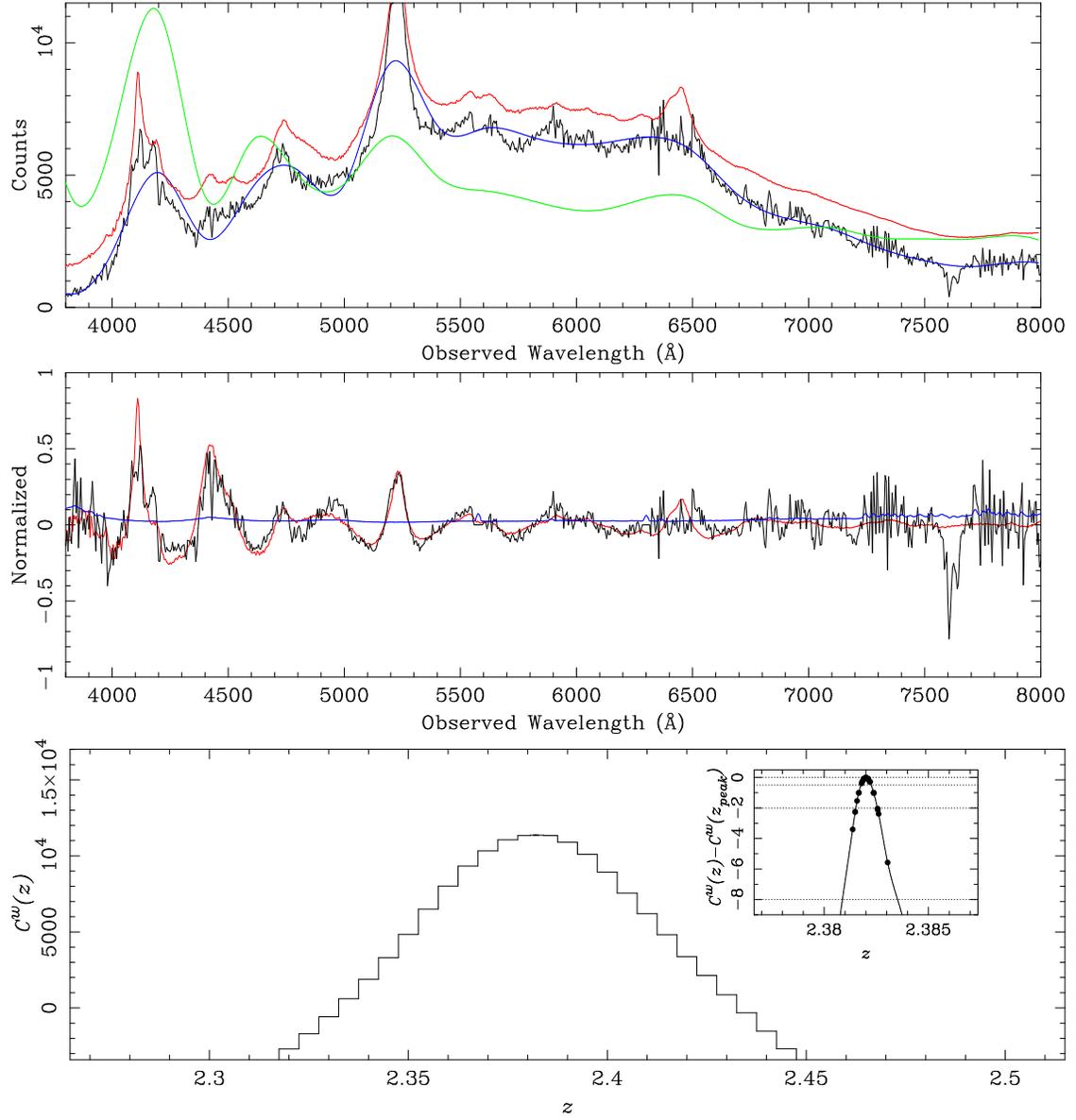


FIG. 9.— Same as Figure 5 for galaxy #821. Note the asymmetric topology of the peak, due to the mismatch between the detailed features of the SDSS composite quasar spectrum (Vanden Berk et al. 2001) and the observed spectrum of #821.

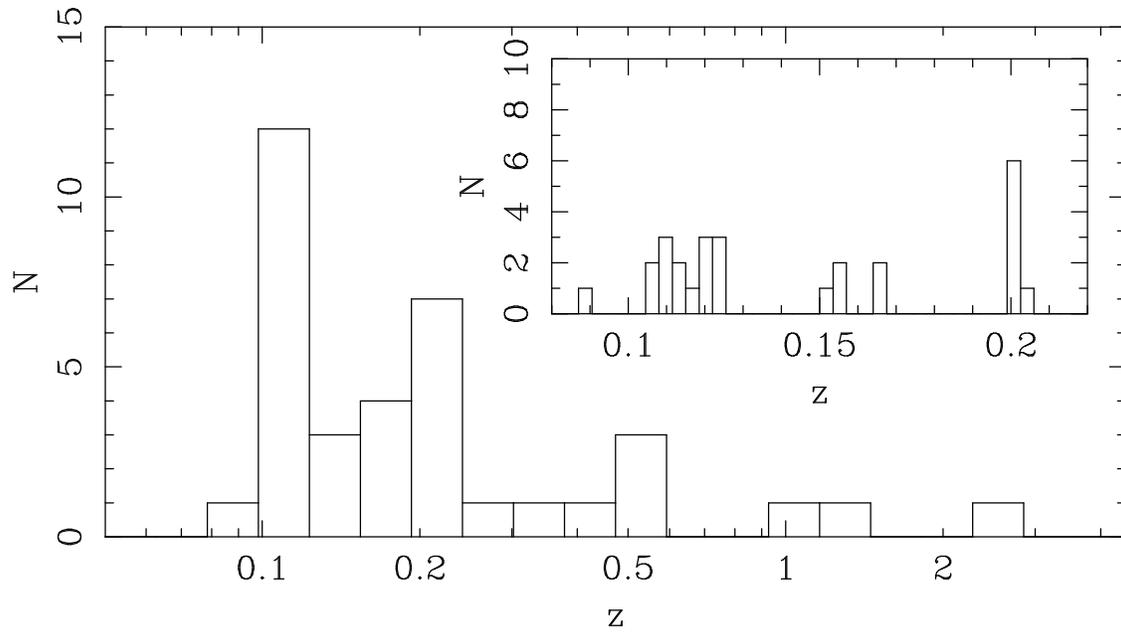


FIG. 10.— The histogram of the 36 redshifts obtained here. The inset shows in greater detail the region about the published redshift of the cluster. The redshift region of the cluster is complicated and does not appear to be a rich gravitationally bound system. The field appears rich because it contains a superposition of four groups spanning redshifts from $z = 0.10$ to $z = 0.20$.

TABLE 1
 REDSHIFT DATA FOR TARGETS IN THE FIELD OF ABELL 2550

ID	R (mag)	α	δ	z	$C^w(z)$	χ_{eff}^2	Redshift Estimator
256	20.1	23:11:46.83	-21:47:09.28	0.20167 (± 9)	220	0.96	Cross correlation
269	18.6	23:11:44.97	-21:46:55.15	0.11229 (± 3)	[OII]3727Å, [NeII]3868Å, [OIII]4363Å H β , [OIII]4959Å, [OI]6300Å
309	19.2	23:11:54.10	-21:46:44.65	0.36888 (± 7)	[OI]6363Å, [NII]6583Å, [SII]6731Å [OII]3727Å, [NeII]3868Å, H β [OIII]4959Å, [OIII]5007Å, H α
343	20.9	23:11:44.10	-21:46:35.82	0.2009 (\pm_1^2)	71	1.47	Cross correlation
352	19.5	23:11:50.03	-21:46:26.24	0.10605 (\pm_7^8)	287	0.93	Cross correlation
391	16.9	23:11:36.42	-21:45:49.82	0.12326 (± 2)	4209	1.78	Cross correlation
422	15.5	23:11:35.75	-21:44:46.18	0.12166 (± 3)	4623	2.72	Cross correlation
453	19.4	23:11:44.23	-21:45:42.19	0.20099 (± 8)	333	1.18	Cross correlation
462	20.1	23:11:31.35	-21:45:41.01	0.20145 (± 8)	[OII]3727Å, H α , [NII]6583Å
471	20.3	23:11:39.40	-21:45:38.08	0.1171 (± 2)	49	1.24	Cross correlation
477	18.0	23:11:40.80	-21:45:19.56	0.12110 (± 4)	990	1.24	Cross correlation
538	20.0	23:11:39.17	-21:45:05.69	0.4222 (± 3)	44	1.23	Cross correlation
102	20.2	23:11:31.25	-21:48:30.90	0.5793 (± 1)	[NeV]3426Å, [OII]3727Å, [NeII]3868Å [OIII]4959Å, [OIII]5007Å
120	18.9	23:11:33.68	-21:48:10.55	0.20434 (± 7)	555	1.83	Cross correlation
165	19.3	23:11:31.09	-21:47:45.51	0.2023 (± 1)	191	1.38	Cross correlation
167	21.3	23:11:26.28	-21:47:51.86	1.415 (± 1)	CIV1550Å, CIII]+SiIII]1900Å, MgII2799Å
176	17.3	23:11:36.82	-21:47:18.54	0.20077 (± 3)	2891	1.64	Cross correlation
259	19.2	23:11:45.73	-21:47:05.18	0.12235 (± 5)	720	1.58	Cross correlation
262	19.5	23:11:44.99	-21:47:04.15	0.16609 (± 7)	506	1.43	Cross correlation
267	21.2	23:11:44.54	-21:47:08.04	1.0608 (± 4)	[OII]3727Å
275	17.4	23:11:31.60	-21:46:43.62	0.12323 (± 5)	[OII]3727Å, H β , [OIII]5007Å H α , [SII]6717Å
335	19.3	23:11:35.67	-21:46:27.78	0.1086 (± 1)	H α , [NII]6583Å
345	18.2	23:11:36.53	-21:46:22.38	0.15475 (\pm_5^5)	46	0.95	Cross correlation
409	19.0	23:11:37.98	-21:46:01.15	0.12103 (± 4)	1187	1.37	Cross correlation
424	16.9	23:11:48.51	-21:45:36.34	0.10572 (± 2)	7945	2.52	Cross correlation
440	18.8	23:11:39.19	-21:45:45.81	0.29362 (± 6)	729	1.13	Cross correlation
554	20.8	23:11:45.55	-21:45:01.46	0.1662 (± 1)	[OIII]5007Å
1193	20.6	23:11:37.72	-21:40:17.16	0.5382 (± 1)	[OII]3727Å
1222	18.6	23:11:34.72	-21:40:36.20	0.1541 (± 1)	204	1.72	Cross correlation
1233	16.7	23:11:44.87	-21:40:27.64	0.11153 (± 6)	[OII]3727Å, [OIII]4363Å, H β [OIII]4959Å, H α , [SII]6717Å
1241	20.4	23:11:42.92	-21:40:19.87	0.1093 (± 1)	46	1.10	Cross correlation
820	18.1	23:11:31.15	-21:41:45.78	0.08824 (\pm_{14}^6)	62	1.36	Cross correlation
821	20.1	23:11:30.49	-21:41:44.39	2.371 (± 1)	OVI1035Å, SiIII1205Å, Ly α 1215Å NV1240Å, CIV1550Å, HeII1640Å NIII]1750Å, CIII]+SiIII]1900Å, CII]2327Å [NeIV]2424Å, [NeIV]2426Å, [OII]2470Å
830	20.5	23:11:49.76	-21:41:06.91	0.1095 (± 2)	46	0.97	Cross correlation
860	23.3	23:11:28.91	-21:41:46.97	0.5373 (± 7)	[OII]3727Å
882	19.1	23:11:36.55	-21:42:03.83	0.1522 (± 1)	[OII]3727Å, H α , [NII]6583Å

Note. — The redshifts listed have not been corrected to the CMB frame. Emission line redshifts were computed after an initial cross-correlation was performed using either a star-forming galaxy spectrum or the Sloan Digital Sky Survey composite quasar spectrum (Kennicutt 1992; Vanden Berk et al. 2001) as templates in the cross-correlation.