Chapter 7

Application to ground-based data

While the tools developed in Chapters 2 & 4 were primarily intended for application to data from space-based transit searches, their application to existing ground-based data provides a fertile testing ground, as well as the potential for scientific return on a shorter timescale. Applications to real data are also bound to highlight problems which did not manifest themselves in the tests on simulated data described so far. In the present Chapter, a brief overview of early experiments with real data is presented. This is not intended as a report on a fully-fledged transit search, but rather as an example illustrating the learning process and the kinds of problems that were encountered.

7.1 The data

The dataset used is part of the University of New South Wales (UNSW) transit search project[1], which uses the Automated Patrol Telescope (APT) at Siding Springs Observatory (SSO). Mike Irwin and myself have been involved in a collaboration with the UNSW transit search team since 2002, our contribution focusing mainly on improving the photometric accuracy of the light curves and on searching for variable stars and transits.

7.1.1 Telescope and instrument

The APT's design is similar to that of a Schmidt camera, but uses a 3-element lens to achieve a wide, corrected field-of-view (FOV). It is entirely computer-controlled and can be operated remotely or – within limits set by the reliability of the weather monitoring systems – automatically. It has a 0.5 m aperture and f/1 optics yielding

[1] See `htpp://www.phys.unsw.edu.au/astro/research/planetsearch.html`.

Figure 7.1: The Automated Patrol Telescope.

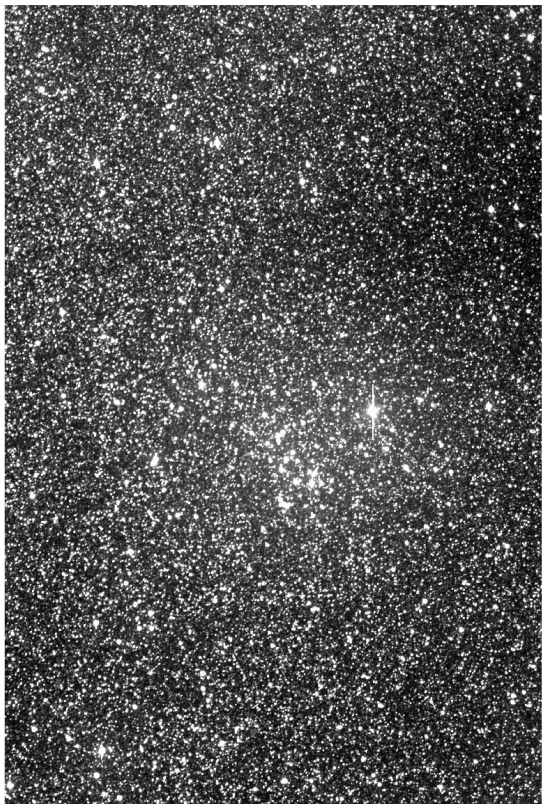

Figure 7.2: A sample APT image of NGC 6633.

a 5-degree flat-field. The current CCD covers a 2×3 degree FOV with 800×1200 pixels, each 9.4 arcsec on a side. Intra-pixel sensitivity changes of the CCD become significant for such under-sampled images, where the light of a star covers only a few pixels. The UNSW team therefore developed an observing technique to overcome this problem, whereby the telescope is systematically moved in a 1×1 or 2×2 pixel 'raster-scan' pattern during each exposure. This has the effect of broadening the PSF slightly, but any intra-pixel sensitivity induced variations are averaged out.

At the time the data used here were taken, the raster-scanning technique was being experimented with and different scan widths were tried on different nights. We used only data taken with a 1×1 pixel scan, which we found to give the best results. The resulting dataset spans 6 nights, with the 4^{th} night (when no scanning was used) missing. Exposure times were 150 s, which leads to stars brighter than $V \sim 9$ being saturated.

The telescope is equipped with broadband *B*, *V*, *R* & *I* filters, but filter changes can only be carried out manually, so that all transit search observations in a given run are performed using a single filter, usually *V* (as is the case for the data used here).

7.1.2 Target

The data used in this Chapter are from a field centred on the open cluster NGC 6633.

NGC 6633 (α = 18^{h}27^m, δ = +06°34′; RA = 11^{h}07m; l = 36.1°, b = 8.3°) is a young open cluster of solar metallicity. Its age is $\log t$ = 8.63, its reddening $E(B-V)$ = 0.18 and its distance modulus $(V-M_V)_0$ = 7.77 (all cluster information taken from Martín et al. 2004). Its youth, moderate reddening and relative proximity, together with the fact that it falls within the COROT visibility zone towards the galactic centre, make it a particularly interesting target for temporal variability studies. Consequently, this cluster was the focus of a multi-site monitoring campaign in August 2002 (Martín et al. 2004). Prior to that campaign, two γ Doradus variable candidates were known in NGC 6633 from monitoring in the Strömgren system (Martín & Rodríguez 2002). Martín et al. (2004) reported on the analysis of data from the multi-site campaign for 192 objects with $V < 12$. Of the 60 stars which were catalogued as belonging to the cluster, they found 12% were variable, and a further 28% were suspected of variability. Variable types included δ Scuti stars, γ Doradus stars, Ap-type variables, eclipsing binaries and long-period variables. The APT data were taken as part of this campaign, but were in the process of being reduced at the time the Martín et al. (2004) paper was written, and were thus not included in that analysis.

Note that, due to the fact that NGC 6633 is situated in the galactic planet, the field is crowded and most of the stars in the magntiude range of interest may not be members of the cluster.

7.1.3 Data reduction and photometry

The data were reduced using a purpose-made pipeline developed by the UNSW planet search collaboration on the basis of the CASU[2] pipeline for the Wide Field Camera (WFC) Survey on the Isaac Newton Telescope (INT). Images were de-biased and trimmed, bad pixels were detected and interpolated over, and a flat field correction was applied using a single master flat-field image generated from twilight flats from several nights.

The photometry and production of light curves were then carried out in Cambridge. A master image was produced by stacking a large number of individual frames, and a deep master catalogue was then generated by running the *imcore* source extraction software on the master image. The source extraction software was also run on the individual frames, allowing a very precise astrometric transformation between each frame and the master image to be computed. Standard aperture photometry (with a core radius of 2 pixels, approximately matching the FWHM of the

[2]Cambridge Astronomical Survey Unit.

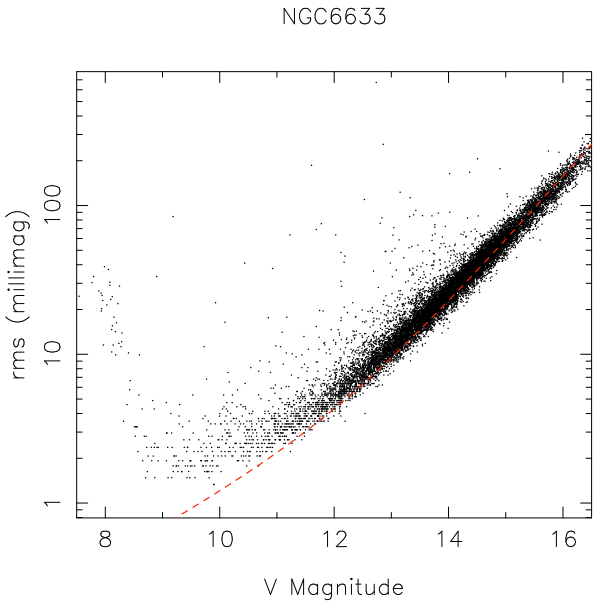

Figure 7.3: RMS precision versus V-band magnitude for 1 night of observations in the field of NGC 6633 (exposure time 150 s, 107 observations). The red dashed line represents the expected rms from object and sky photon noise.

images) was then carried out on each frame using the positions derived from the master catalogue. Finally, a correction equal to the average departure from the light curve median over each frame was applied (this is equivalent to a very crude relative zero point, including atmospheric extinction correction). Frames for which the global rms exceeded 10 mmag were discarded as indicative of non-photometric conditions.

The resulting rms precision is ~ 2 mmag for unsaturated stars brighter than $V \sim 10.5$ within a single night (see Figure 7.3). The backbone of points at the bright end is shifted slightly upward from the expected value from photon noise statistics, showing that other low-level noise sources remain.

7.1.4 Post-processing

The light curves contain residual systematic trends, which result in a degradation of the precision when several nights are joined together (see Figure 7.4, left panel). Examination of the light curves reveals that most display parabolic trends within each night (as one would expect from atmospheric extinction and airmass dependent seeing), as well as night-to-night shifts. These systematics are very damaging for transit searches, as downwards trends at the edges of nights mimic transit ingress or egress events, while a long term trend or night-to-night shift can lead to a spuriously high transit detection statistic because the duration of a night of observations (≤ 6 hours in this case) falls within the range of reasonable transit durations.

A more accurate extinction correction, by which nightly extinction coefficients are derived from iterative parabolic fits to the light curves, was initially experimented

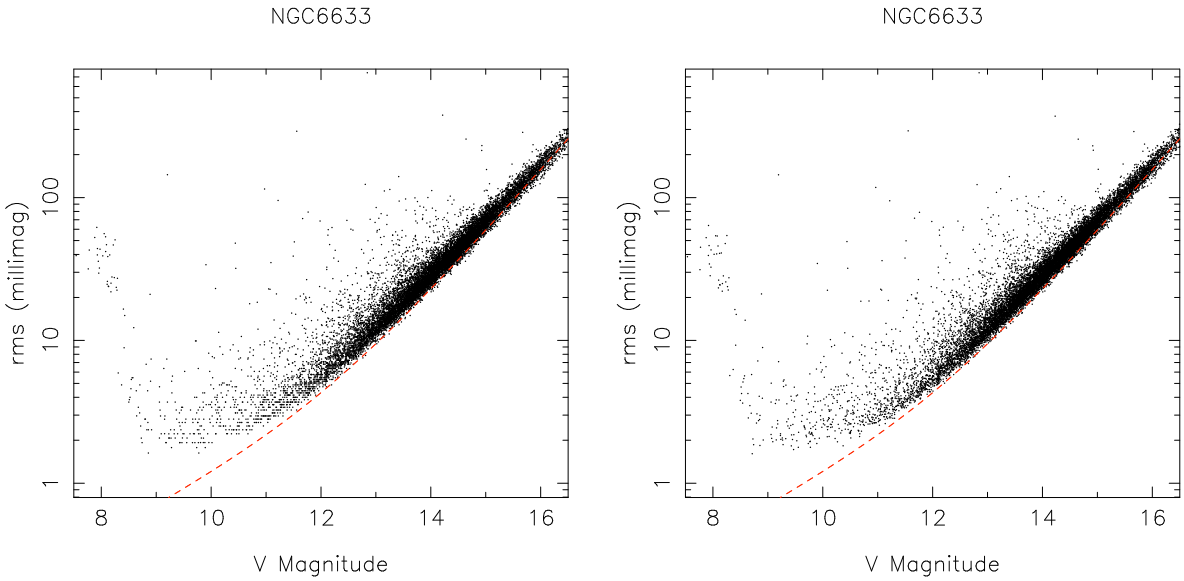

Figure 7.4: RMS precision versus V-band magnitude for the 5 nights of observations in the field of Open Cluster NGC 6633 (exposure time 150s, 417 observations) without (left) and with (right) systematics correction.

with, but in the present case was not found to significantly improve the results.

The standard practice to obtain a "systematics free" light curve for a given object is to identify trends common to a number of nearby comparison stars, which are assumed to be constant apart from noise and the systematics of interest, and to subtract those trends from the light curve of the star of interest (see for example Gilmore 1979). In the age of small field-of-view detectors with non-linear response, care had to be taken to account for the non-linearities and to ensure that the comparison stars really were constant, or to account for the possibility that they might not be. With wide field-of-view CCDs, many comparison stars are available. One can rely on averaging over many comparison stars to iron out the effects of intrinsic variability of any given comparison star. This is, in effect, implemented through the global zero-point correction described at the end of Section 7.1.3.

However, one must also take into account the possibility that the systematics are position dependent, as differential airmass and PSF variations across the field-of-view may become important. In the context of an INT/WFC monitoring campaign in the field of open cluster NGC 6819, Street (2002) implemented a position-dependent zero-point correction. The zero point for each frame (i.e. the median, for a given image i, of the individual deviations $\delta_{ij} = \overline{m_j} - m_{ij}$ from the light curve average or median $\overline{m_j}$ for each star j) is computed not for the entire image but for a square region of the image (the optimal size of which is determined through trial and error).

In the present work, we implemented a similar procedure, but rather than subtracting the zero-point for each region we proceeded, once the zero-points were

computed for each region of an image, to fit a 2–D polynomial to those zero-points as a function of x and y position. The correction applied for each star in each image was then derived from this fit, rather than from the zero-points themselves. It is therefore a smooth function derived from the entire grid of regions rather than a discontinous set of values. Some care must be taken in the calculation of the zero-points for each region. The size of the regions was set to 100×100 pixels, after some experimentation with different values. This is low compared to the value of 500×500 used by Street (2002), but the APT pixels are much larger than the INT/WFC pixels. The smooth nature of the correction also makes it less sensitive to the size of region used. Only objects classified by the source extraction software as stellar (i.e. circular) and falling within the magnitude range $9 \leq V \leq 16$ are included in the zero-point calculation. Each star is weighted according to the inverse of the overall variance of its light curve. Cases where $|\delta m_{ij}| > 0.25$ are also excluded, as they correspond either to true variable stars or to outlying points in an otherwise constant light curve. A 2^{nd} order polynomial was used. The fits are not improved by increasing the order of the polynomial.

This method was found to yield a small improvement in the overall rms (see Figure 7.4, right panel), reducing the spread about the backbone of points for a given magnitude as well as the number of stars with rms values significantly above the expected value for their magnitude. It made a significant difference to the number of spurious transit candidates with high detection statistics.

While the steps described here should remove any position dependent zero-point effects, it seems that not all the systematic trends are fully removed. Certain groups of light curves, unrelated in position or magnitude, still appear to show common trends. Stars of potentially very different temperature and colour are included in the computation of the correction applied to any given light curves, and this may affect the results, but no colour information is available to – for example – select colour-dependent comparison stars. Other, unforeseen parameters may also affect the systematics. Street (2002), who also observed such residual systematics after applying her position-dependent zero-point correction, implemented a method to identify those light curves whose residuals were most similar to each light curve in turn, and to derive and apply a correction based on that set of light curves. A similar approach has recently been published by Kovács et al. (2004), while a PCA-based approach for the same purpose has been proposed by Tamuz et al. (2005). These methods, or variants thereof, will be investigated in the future to see if they can further improve the quality of the light curves.

After the systematics correction, data points deviating from the light curve median by 5σ or more (σ being the MAD-estimated scatter rather the than the standard

deviation) were removed from each light curve. To avoid removing real features, this clipping stage was suppressed in light curves with more than 5 such outlyers (which are generally eclipsing binaries or flare stars).

7.2 Transit search

Before entering the transit search itself, an aside on variable star searching is appropriate. A search for variable stars, using the same reduced chi-squared criterion and sine-fitting technique as mentioned in Section 6.3.1.3 for the COROT blind exercise, was carried out on the 5 nights of APT data, though without the post-processing steps described in the previous Section, which had not yet been developed. This variable star search, some of the results of which were published in Hidas et al. (2003), was intended primarily to demonstrate the precision achieved by the APT observations and the potential of the survey, and concerned only the thousand brightest stars. We identified approximately 100 variables, including all those reported by Martín et al. (2004) in the magnitude range covered by the APT data, but also tens of new variables of various types, from low-amplitude irregular variables to eclipsing binaries, multi-periodic pulsating stars and RR Lyrae stars. Variable star identification and classification is an integral part of the scientific exploitation of transit search data, and efforts in this direction are ongoing, but we concentrate in the present Chapter on the transit search itself.

The transit search algorithm described in Section 2.2, modified to search over a range of trial durations as well as periods and epochs, was applied to all light curves. Trial durations ranged from 0.04 to 0.25 days (1 to 6 hours) and 0.8 to 4.0 days, with increments of 0.01 day for both. The trial epoch increment was set to the average interval between consecutive observations (excluding daily gaps), which was 0.0026 days (3.75 min). The parameters of the best transit candidate (corresponding to the highest multiple transit detection statistic S) were recorded for each light curve.

In order to establish an appropriate detection threshold, the cumulative distribution of the multiple transit detection statistic S obtained for each light curve was plotted. The result is shown in Figure 7.5 (black line). It is immediately apparent that the vast majority of the light curves contain at least one event with $S \geq 6$, the threshold derived from Monte Carlo simulations to be optimal for white noise only (see Section 2.2). While there is a distinct tail above $S > 30$, these cases correspond to very deep eclipses ($> 10\%$) which are unlikely to be of planetary origin, and restricting oneself to those would defeat the point of the transit search. There is no clear feature in the distribution to indicate an appropriate cutoff lower down, and many of the cases in the range $10 < S < 30$ correspond to spurious detections due to

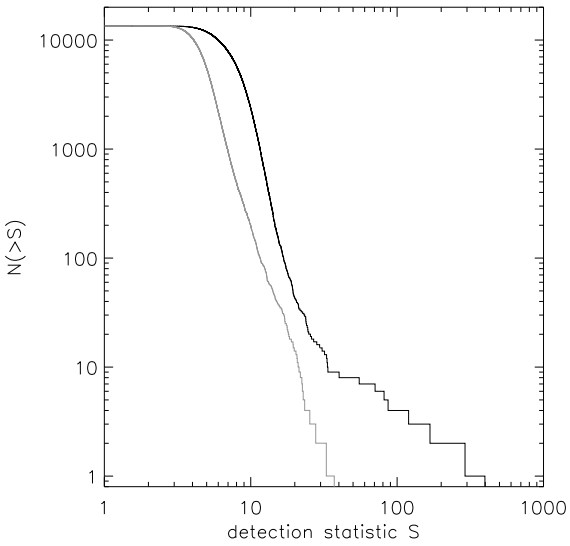

Figure 7.5: Cumulative distribution of the multiple transit detection statistics for all stars with $9 < V < 16$. Black line: no filtering; gray line: after filtering of the long term trends.

night-to-night shifts or long-term trends in the light curves.

Given the lack of a natural way to determine an optimal threshold, one can select one that will give the desired number of candidates. Artificial though it may be, this provides a means of making a rapid foray into the data, and the threshold can be relaxed later if wished. But first, it is desirable to try to weed out at least some of the spurious detections automatically. To this end, two approaches were investigated.

The first was to repeat the transit search, preceded by a long-term variability filtering stage. The filter used was a simplified version of the least-squares filter described in Chapter 4, where rather than fitting sine curves up to relatively high frequencies and applying a matched or Wiener filter, one simply fits and subtracts sine-curves up to a given frequency. In effect a highpass filter, this is almost identical to the approach used by team 3 in the COROT blind exercise, though here sine curves of increasing frequency were fitted and subtracted one by one rather than simultaneously. The fitted frequencies were harmonics of the fundamental frequency $1/2T$ (T being the total time span of the light curves), and the highest fitted frequency was chosen to correspond to five times the longest trial duration (i.e. 1.25 day, which ensures the removal of any long term trends and most night-to-night shifts). Note that the dataset is very short, hence only a total of 8 frequencies are fitted in this way.

The resulting detection statistic distribution is shown as the gray line on Figure 7.5. The tail of high S events has disappeared (as the filter removes some signal from very deep eclipses), and the overall distribution has been shifted to lower S. Examination of the 100 light curves with the highest S (which brings us, in this case, down to $S = 10.2$) reveals that long term trends or night-to-night shifts are no longer a

problem, but many spurious detections arise at periods close to 1.0, 1.5 and 2.0 days. These are due to downward trends at the edges of nights, which are still present in many light curves despite the post-processing steps described in Section 7.1.4. We find $\sim$ 30% of the light curves show transit or eclipse-like dips, and $\sim$ 10% show other forms of variability (RR Lyrae-like, short-period pulsations, ...). The remaining $\sim$ 60% either appear to be spurious detections due to night-edge effects, or have insufficient phase coverage to establish confidence in the detections. Excluding periods within 0.02 day of 1.0, 1.5 or 2.0 days, removes two of the shallower eclipse-like dips and one variable while reducing the number of spurious detections to $\sim$ 10.

Another approach is to examine the parameters of the high S detections in the case where no filter was used, to see if a criterion based on these parameters could separate the true transit-like events from the detections induced by long-term trends or night-to-night shifts. Of the cases with $S > 20$, it is notable that all but one of the spurious detections have durations larger than 0.23 day, while all the transit-like events have durations below 0.17 day. The light curves for the 100 highest S cases with $d < 0.17$ day were thus examined by eye. We found $\sim$ 40% had transit-like eclipses, 3 light curves showing signs of other variability, 2 due to night-to-night shifts, and the remainder could either be ascribed to night-edge effects, or the phase-coverage of the transit candidate was too poor to assess it.

It is interesting to note that, whether using filtering or not, a number of shallow transit candidates fall at the end of the first night. The fact that many stars simultaneously show a dip at the same time suggests that it is a systematic effect. Such 'transits' are also only partially covered, the egress falling after the end of the night. Shallow transits at the end of a night were thus only accepted if more than one transit was seen and the phase coverage of the transit was sufficient.

A final list of 40 transit- or eclipse-like events was compiled by merging the candidates from the two approaches described above that appeared convincing upon visual inspection of their light curves. The resulting list is shown in Table 7.1. Most candidates were detected through both approaches, but 4 were detected with the first approach (filtering) only and 12 with the second approach (no filter) only. Plots of the phase folded light curves for these candidates, as well as examples of some typical spurious detections, are shown in an Appendix at the end of this Chapter.

7.3 Discussion

The transit search tools developed in Chapters 2 & 4 were applied to a typical ground based transit search dataset, and allowed the detection of several tens of transit-like events, 7 of which have depths $\leq$ 50 mmag. The light curve analysis was carried out

Table 7.1: Transit or eclipse candidates with their parameters.

ID (1)	V (2)	σ (3)	d (4)	p (5)	e (6)	S_M (7)	dV (8)	dF/F (9)	N_{tr} (10)	approach (11)	(12)
164	9.963	0.00268	0.123	0.879	0.883	397.880	151	13.0	1	1	b
355	10.875	0.00548	0.123	1.778	1.095	77.712	55	5.0	2	1	b
742	11.675	0.00712	0.071	1.089	1.589	19.670	21	1.9	2	2	2
1007	11.999	0.00877	0.091	0.995	1.076	21.171	18	1.7	5	1	b
1162	12.119	0.01507	0.060	1.024	1.810	88.278	198	16.7	3	1	b
1397	12.295	0.01437	0.144	0.858	1.549	151.343	259	21.3	2	1	b
1640	12.461	0.06844	0.060	0.966	1.131	16.621	126	11.0	4	1	b
1752	12.524	0.01669	0.060	1.106	0.844	18.052	40	3.7	3	1	b
1816	12.568	0.01746	0.154	1.118	1.757	184.181	410	31.5	1	1	2
1965	12.646	0.01015	0.112	1.109	1.773	27.842	41	3.7	1	1	b
2041	12.683	0.01174	0.070	0.879	1.643	28.836	64	5.8	1	1	b
2042	12.682	0.01960	0.164	0.811	0.938	12.821	22	2.2	3	1	2
2272	12.777	0.01104	0.123	0.801	1.410	85.210	131	11.4	1	1	b
2547	12.895	0.06896	0.112	1.309	0.886	15.576	139	12.1	2	1	b
2787	12.983	0.02092	0.133	0.879	0.881	63.703	183	15.6	1	1	b
2829	13.001	0.02622	0.091	0.971	1.089	13.888	30	2.7	5	1	b
2856	13.008	0.04935	0.081	1.008	0.857	14.577	69	6.2	4	1	2
2893	13.030	0.01751	0.050	1.233	0.844	25.226	90	8.0	2	1	b
3498	13.217	0.03943	0.133	1.109	1.778	21.229	112	9.8	1	1	2
3557	13.244	0.09491	0.091	2.061	0.920	11.738	116	10.1	3	2	1
4046	13.382	0.01630	0.112	1.096	1.799	12.242	28	2.6	1	1	2
4095	13.392	0.11796	0.070	1.895	1.058	11.380	164	14.1	3	2	1
4254	13.435	0.01525	0.102	1.212	1.984	13.682	32	2.9	1	2	1
4290	13.440	0.03046	0.133	1.217	0.969	46.500	160	13.8	2	1	b
4389	13.470	0.02209	0.155	0.834	1.120	24.535	55	5.0	3	1	2
4585	13.521	0.04540	0.174	1.075	1.637	33.823	138	12.0	3	2	1
5198	13.658	0.02101	0.112	0.869	0.961	33.480	101	8.9	1	1	b
5355	13.694	0.07742	0.070	1.930	1.016	14.462	131	11.4	3	1	b
5638	13.751	0.03076	0.060	0.820	0.961	16.576	71	6.3	3	1	b
5782	13.781	0.04338	0.124	1.080	0.957	33.349	156	13.4	3	1	b
6340	13.911	0.08166	0.112	0.826	1.246	20.595	200	16.8	2	1	2
6897	14.021	0.10043	0.102	0.997	1.063	21.349	199	16.8	5	1	2
7249	14.095	0.03491	0.081	1.136	0.925	13.715	62	5.6	2	1	2
7295	14.101	0.05452	0.071	1.200	0.878	22.212	182	15.5	2	1	b
7752	14.201	0.05209	0.092	1.049	1.781	13.291	73	6.7	3	1	2
8208	14.291	0.09715	0.081	1.372	1.906	33.079	476	35.5	2	1	b
8209	14.290	0.04938	0.060	1.464	1.951	24.600	203	17.1	2	1	b
9963	14.632	0.07760	0.071	0.988	0.875	21.218	241	20.0	3	1	b
11231	14.918	0.22918	0.050	0.978	1.001	18.914	464	34.8	5	1	b
11566	14.997	0.10667	0.060	1.824	1.063	12.359	196	16.5	2	1	2

Columns: (1) star ID; (2) median magnitude; (3) fractional light curve scatter; (4) transit duration (days); (5) transit period (days); (6) time of start of first transit (days, first data point is at t = 0.88 day); (7) detections statistic; (8) transit depth (mmag); (9) fractional transit depth in flux units (%); (10) number of transits actually sampled; (11) approach the parameters were obtained with; (12) approach the event was detected with (b means both).

in just over one working week, i.e. the time span of the observations. This bodes well for the prospects for semi-automated transit searches in such datasets, but a number of issues remain to be addressed.

The UNSW team have independently carried out a transit search in this same dataset, though using all eight nights collected in 2002 and not just the nights with 1×1 rastering (Hidas et al. 2005). Although they did employ a systematics removal technique similar to the 2–D polynomial fitting described in Section 7.1.4, they still found too many false alarms were induced by nightly trends in the data to be able to use automated transit search algorithms, and therefore opted for a visual inspection of each light curve, limiting themselves to the 2–3000 brightest stars. They thus identified 4 transit candidates with depths below 100 mmag, in light curves 306, 742, 2041 and 1965, as well as two lower confidence candidates, in light curves 795 & 252. 742, 2041 and 1965 are amongst our candidates, but 306 was not included because it only gave rise to a detection statistic of $S \sim 8$, below the cut off in both candidate selection approaches we investigated. The UNSW went on to follow-up their 4 candidates, first checking the transit repeated in data from subsequent seasons, then examining Digitized Sky Survey plates to establish the degree of blending in the APT apertures, and finally obtaining higher resolution time-resolved photometry and low resolution spectroscopy. They found the brightest star in the apertures corresponding to light curves 306, 2041 and 1965 to show large amplitude radial velocity variations typical of stellar mass eclipsing binary objects. In the case of light curve 742, the transit occurs across the second brightest star in the APT aperture, and spectroscopic observations showing a double lined profile whose separation varies in phase with the transits suggest a system composed of two nearly identical K7 stars in a grazing eclipsing binary configuration.

Whether my implementation of the post-processing steps was more successful, or whether I simply accepted a need to weed through a larger number of false alarms, I was able to apply a mostly automated transit search tool to these data. This allowed me to search all the light curves and not only those of the brightest 2–3000 stars (though of course the chances of detecting a planetary transit are much lower in the light curves of fainter stars). I identified, in addition to 3 of the candidates found by Hidas et al. (2005), a number of shallow transit candidates (light curves 1007, 1752, 2042, 2829, 4046 and 4254). Light curve 1007 has relatively poor phase coverage and may have been rejected by the UNSW team on that basis, though it has been included here because there do seem to be a few points belonging to the transit egress. It will be very interesting to see whether the transits repeat at the predicted times in the 2003 and 2004 data. Light curve 1752 shows clear signs of out-of-eclipse variations, and the eclipses are thus unlikely to be due to a

planet, but the light curve shape is intriguing. Although the period identified for light curve 2042 is clearly incorrect, one can see from the phase-folded light curve that it is an eclipsing binary with visible secondary eclipses. Light curve 2829 is an interesting case with a clean transit which may be flat bottomed, and is certainly worthy of follow-up. The reliability of cases 4046 and 4254, which are much fainter and have larger rms, is harder to establish, as only one transit was observed in each case. Should additional transits be found in the 2003 and 2004 data, these would also make interesting candidates.

Hidas et al. (2005) also found 9 deep eclipsing binary (W Uma and Algol) systems and 31 other variable stars. We identified 6 of the EBs (light curves 1162, 1640, 2272, 2547, 2787 & 2856). The other 3 were all W Uma type systems which are almost continually in eclipse. For those, the assumptions (on which the transit search algorithm is based) that the mean out-of-transit level is equal to the overall light curve mean, and that the time in-transit is short compared to the total light curve duration, break down. In practice, the eclipse minima are detected but the detection statistic is below the threshold used here ($S \sim 10$ in all three cases), because the light curve scatter is high. Failure to detect these systems is not considered as a failure of the algorithm, as it was not designed specifically to detect such signals.

However, the 3 eclipsing binaries and the one transit candidate of Hidas et al. (2005) that were not included were only slightly below our detection threshold. This highlights the fact that, given the current state of the post-processing and transit search process, examining the 100 light curves with the highest S is likely to miss many interesting candidates. Ideally, one would wish to examine all light curves with $S > 6$, but in the present case that would include 1000 cases, the vast majority of which are likely to be false alarms. Additional work must therefore concentrate on reducing the false alarm rate before a relatively complete automated transit search can be carried out.

Three principal sources of false alarms[3] were encountered: long term trends, night-to-night shifts and night-edge effects (generally at the end, rather than the beginning, of each night). The first two are effectively removed by a simple high-pass filter implemented via sine-curve fitting, but at the cost of removing part of real transit-like signals and thus a loss of sensitivity. The third suggests that the systematics correction applied can be improved. Diagnostic tests to understand the origin of the uncorrected systematics are underway. There appears to be a strong gradient in the summed light curve residuals as a function of y-position at one edge of the

[3]By false alarms, we mean here detections of events that turn out not to be transit-like, as opposed to detections of transit-like events that turn out not to be planetary (e.g. eclipsing binaries). Short period variable stars, which sometimes give rise to high transit detection statistics, are also not considered as "false alarms" here because they are indeed variable in a way which *should* trigger the algorithm.

CCD (this could be due, for example, to detector artifacts or vignetting), which is not completely removed by the systematics correction applied. A number of potential improvements to the systematics correction are under investigation, but the most effective may be to exclude stars falling in the problematic part of the CCD. Alternative systematics correction methods from the literature are also under review.

It is important to note that, while some degree of visual examination of candidates' light curves will always be necessary, the goal of any automated transit search system is to minimise this need. The amount of visual inspection carried out in the present Chapter – where $\sim$ 60% of the light curves examined turned out to be spurious candidates – is certainly suboptimal.

The dataset investigated here was very short, and some of the problems encountered in this analysis can be expected to be less important for longer datasets. In particular, sensitivity to spurious periods close to one day or aliases thereof should decrease. NGC 6633 was observed during several runs in 2003 and 2004 with the APT, and the post-processing of these data is underway. If any of the transit candidates which are listed here remain after the analysis of the data from the 2004 season, follow-up observations will be undertaken, starting with higher resolution multi-band photometry to identify possible blends and provide a preliminary assessment of cluster membership.

Appendix: Phase-folded light curve plots

Figures 7.6 to 7.13 contain plots of the normalised phase folded light curves for all the transit or eclipse candidates. There is one row per transit candidate. In each case the left panel shows the full phase-folded light curve, the right panel the phase folded light curve around the transit. The red line in both panels shows the box-shaped transit model corresponding to the reported parameters. The light curve ID is indicated in the top of the left panel. The transit duration, period and epoch (in days) are listed at the bottom of the left panel, the transit depth (in mmag and in %) at the bottom of the right panel. Note that the unfiltered light curves are plotted in all cases, even when the detection was made after filtering, as the filter sometimes distorts the transit shape. In several cases (e.g. ID 2547) the detected period is clearly incorrect, but there is a transit-like event nonetheless. In such cases, the phase-folded light curve is plotted with the detected period, as the plots are automatically produced by the program implementing the detection algorithm.

Figures 7.14 to 7.16 contain similar plots for some representative spurious detection cases (long term variations, night-to-night shift and night-edge effects respectively). Those figures contain an additional panel with the unfolded light curve, which

helps highlight the spurious nature of those candidates.

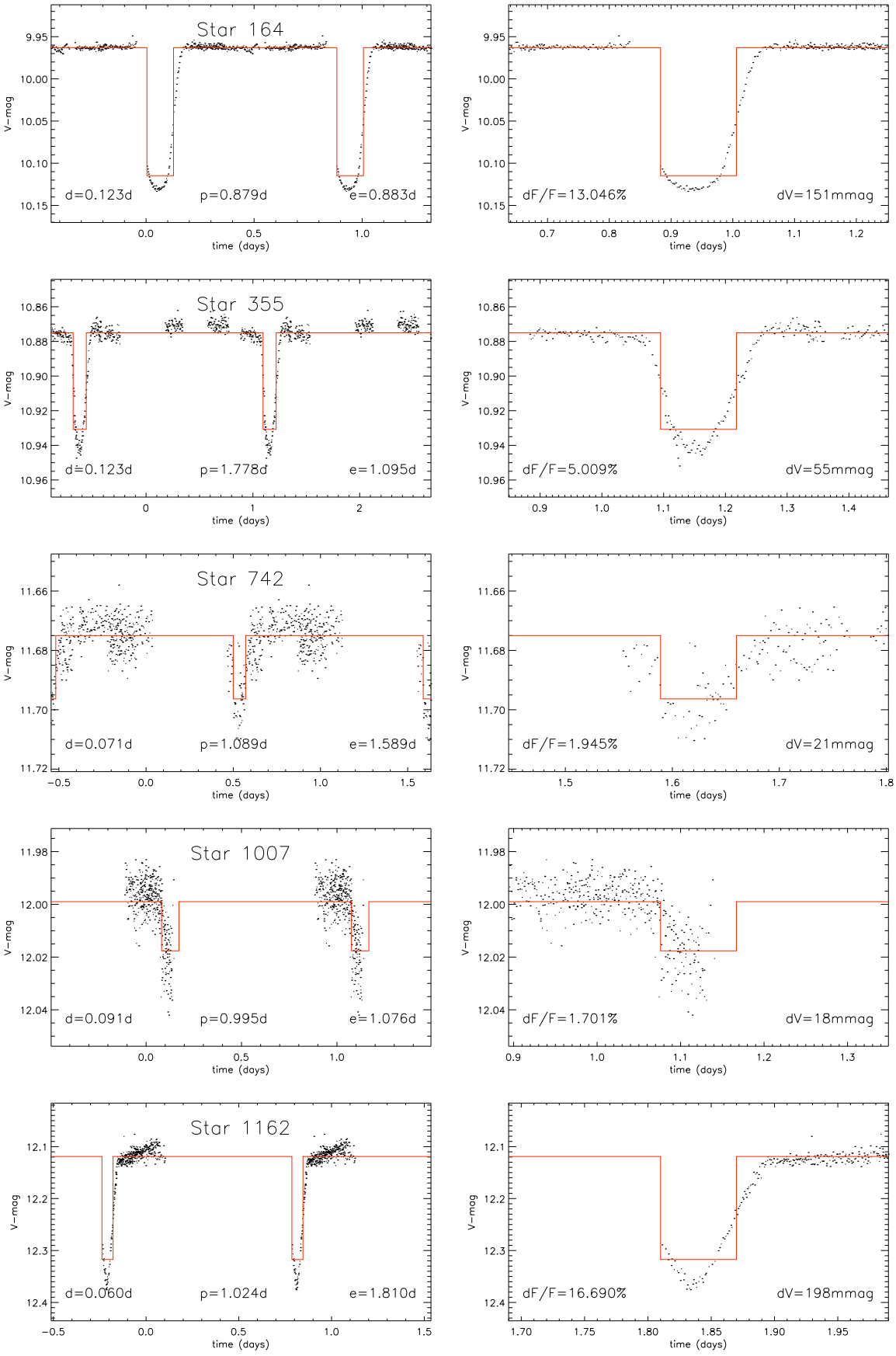

Figure 7.6: Light curves of transit / eclipse candidates – page 1.

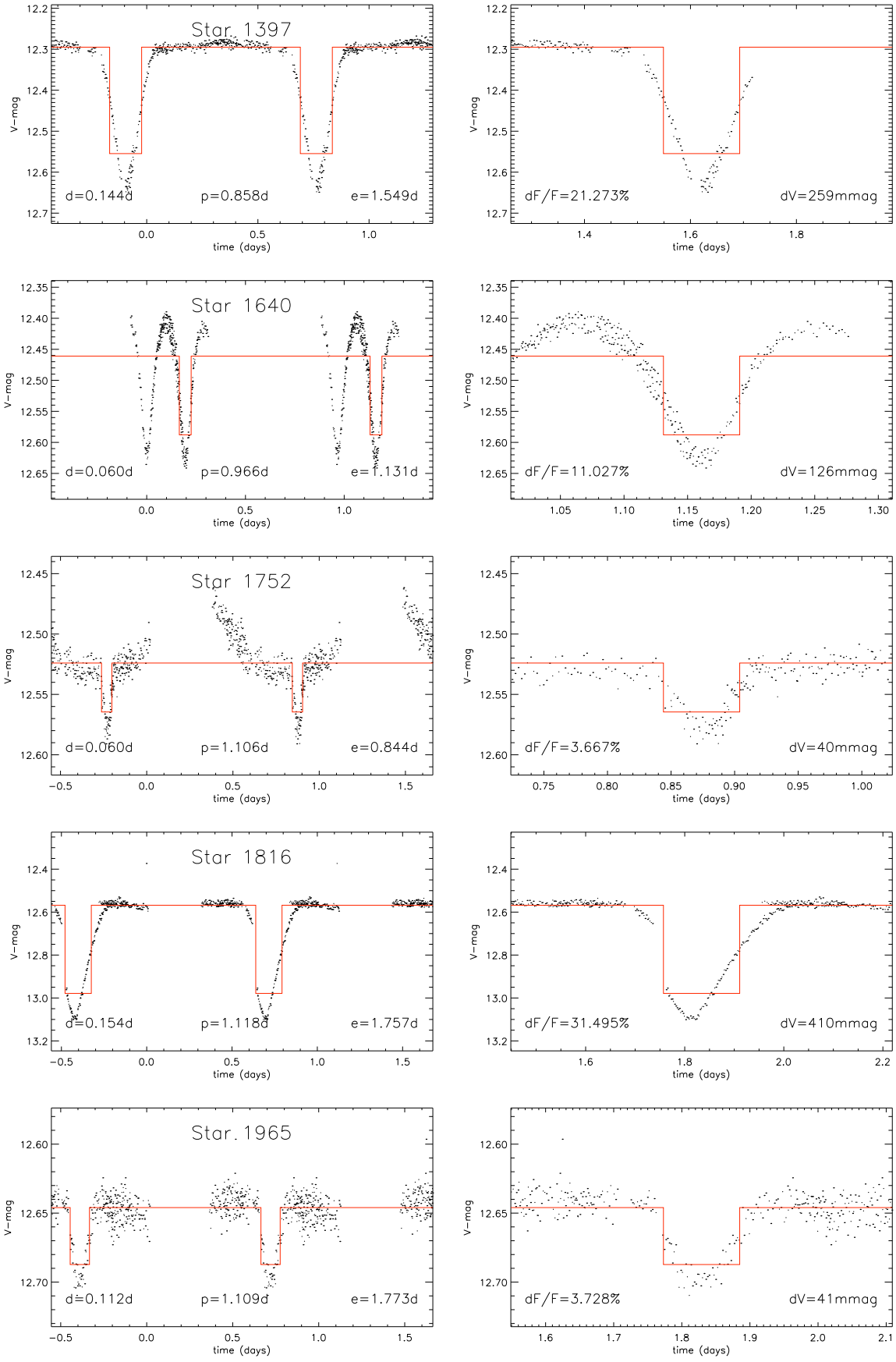

Figure 7.7: Light curves of transit / eclipse candidates – page 2.

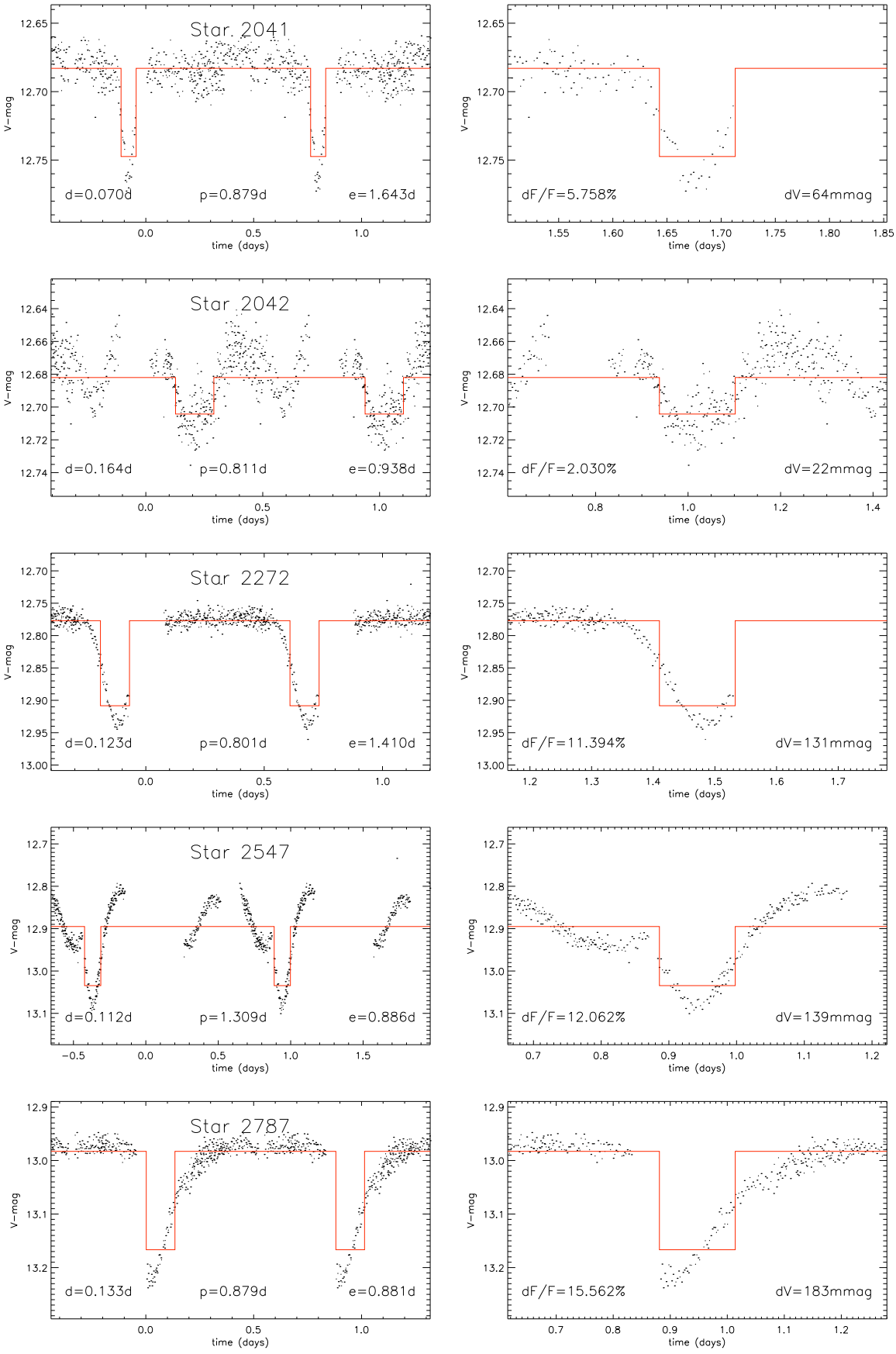

Figure 7.8: Light curves of transit / eclipse candidates – page 3.

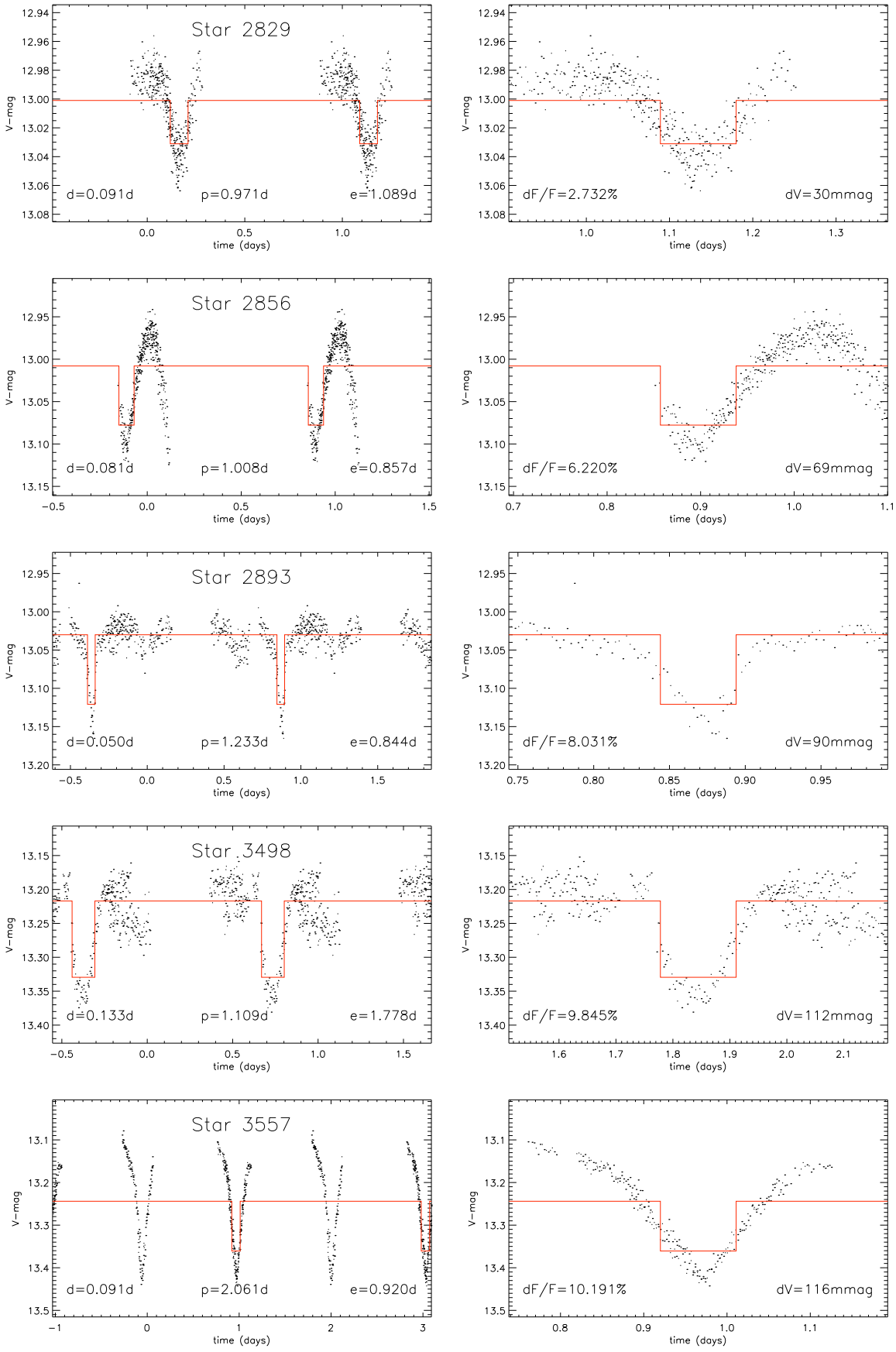

Figure 7.9: Light curves of transit / eclipse candidates – page 4.

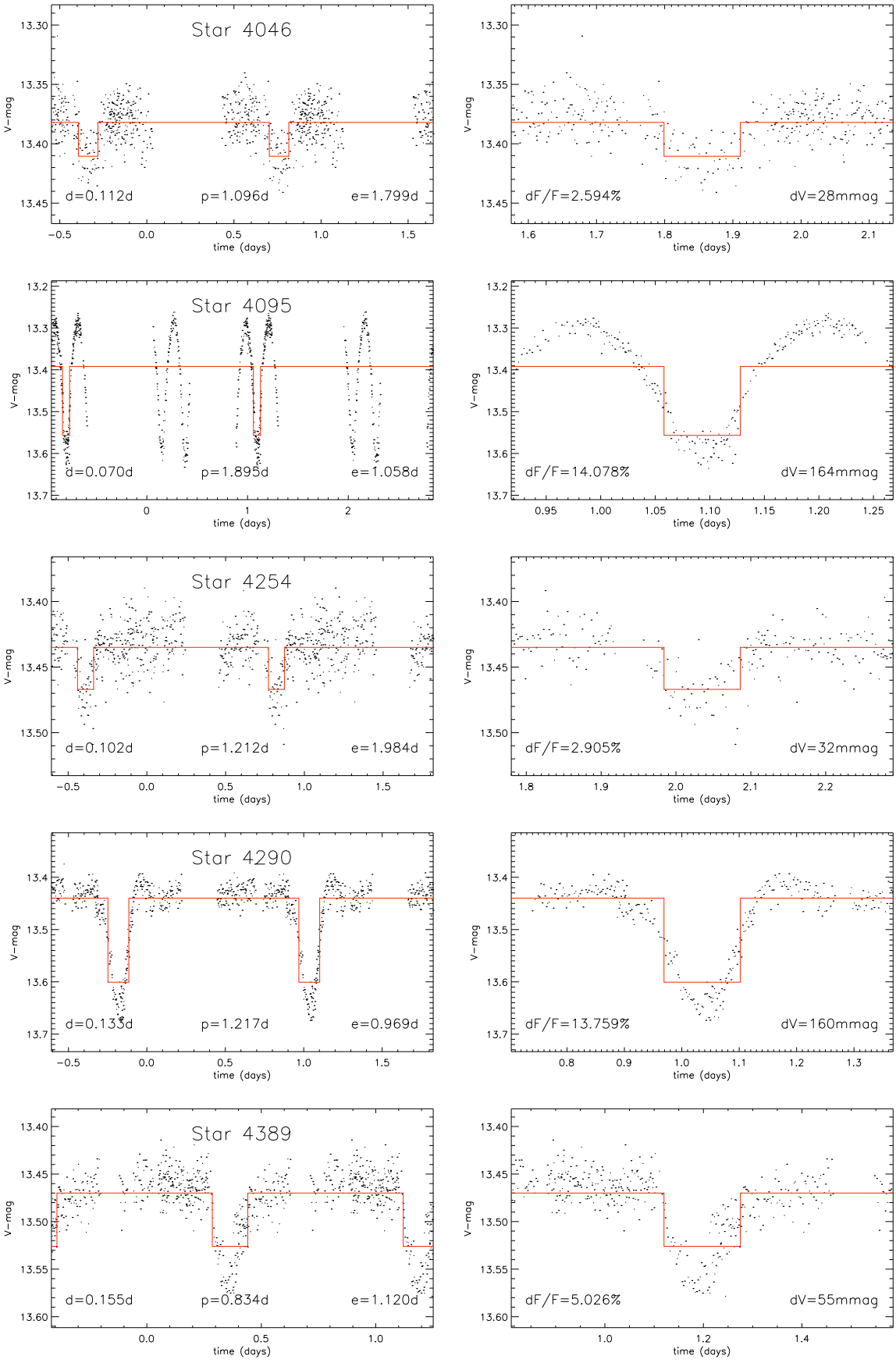

Figure 7.10: Light curves of transit / eclipse candidates – page 5.

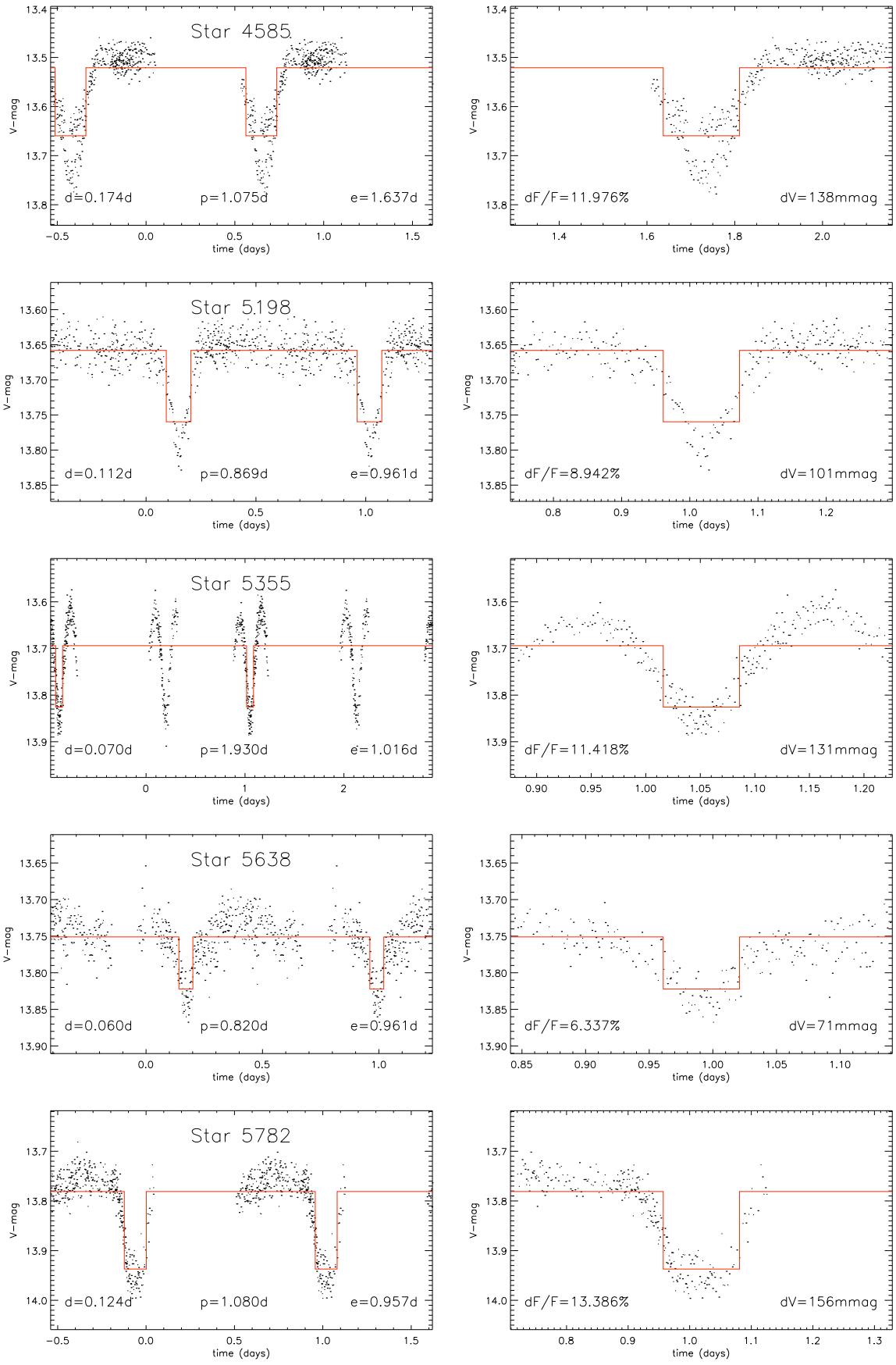

Figure 7.11: Light curves of transit / eclipse candidates – page 6.

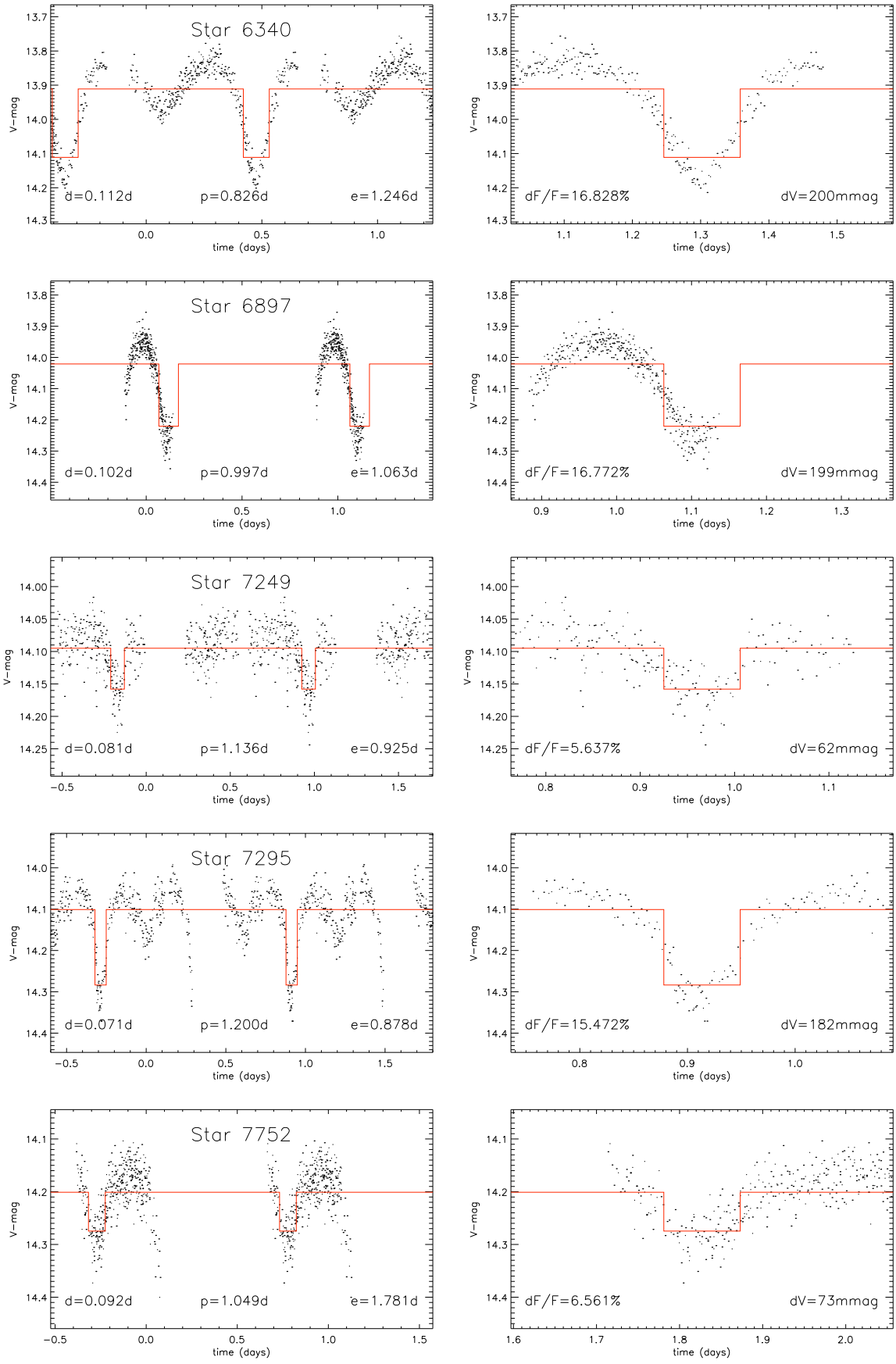

Figure 7.12: Light curves of transit / eclipse candidates – page 7.

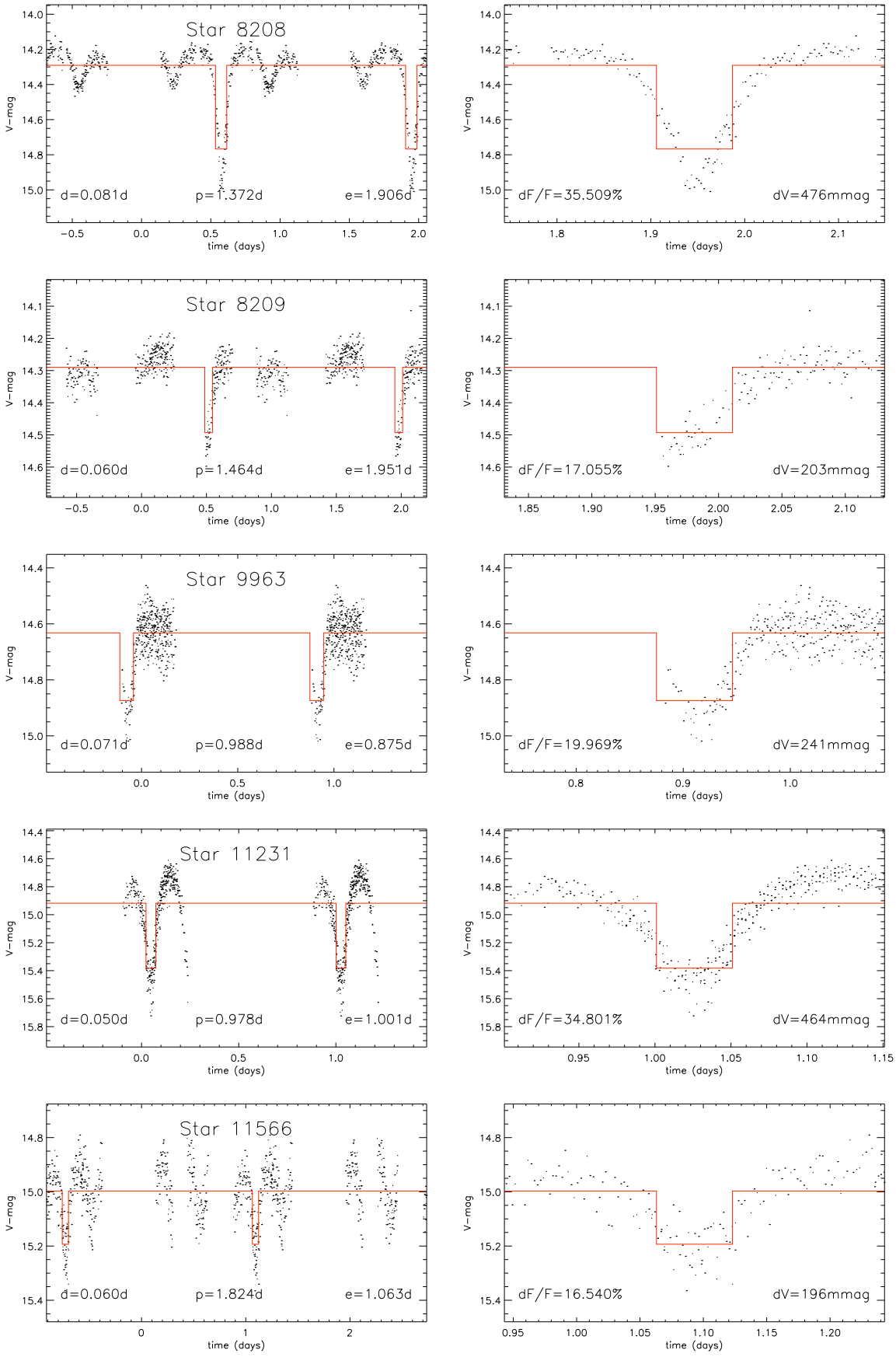

Figure 7.13: Light curves of transit / eclipse candidates – page 8.

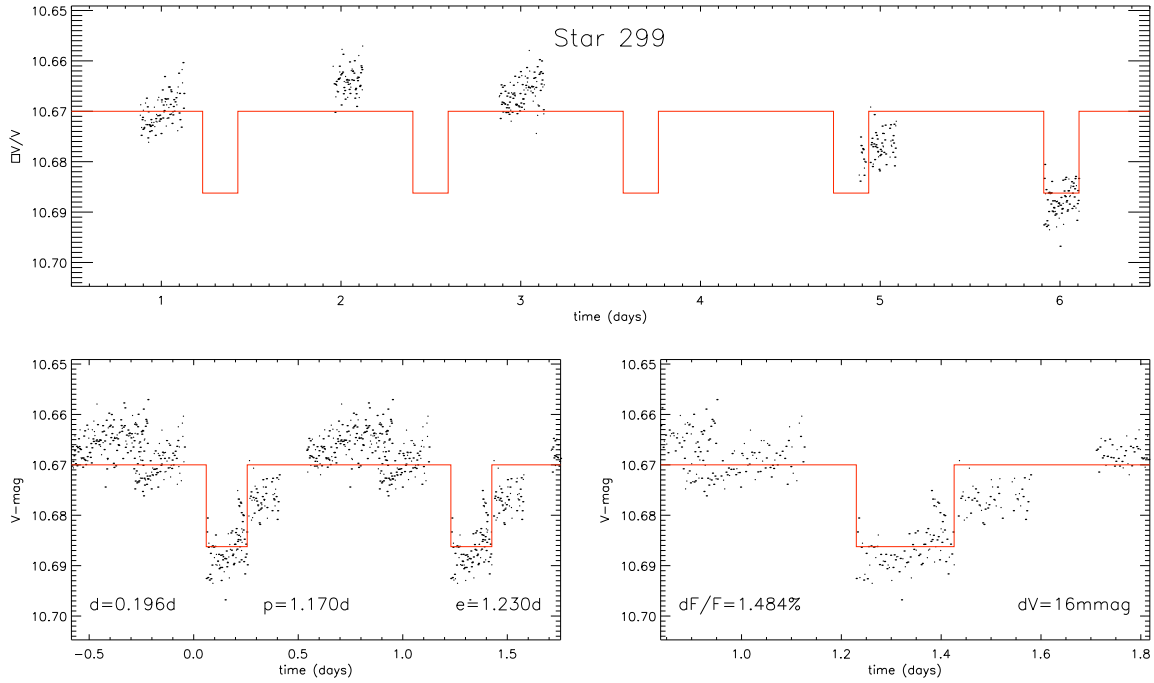

Figure 7.14: Spurious candidate example: long-term variations.

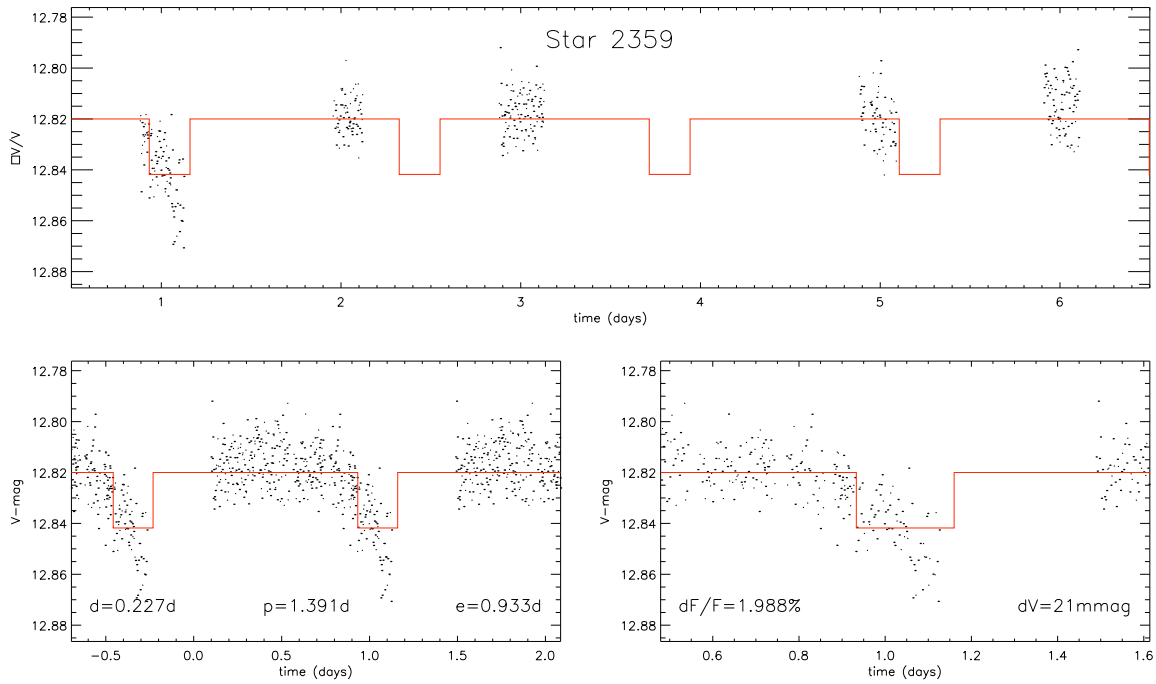

Figure 7.15: Spurious candidate example: night-to-night shifts.

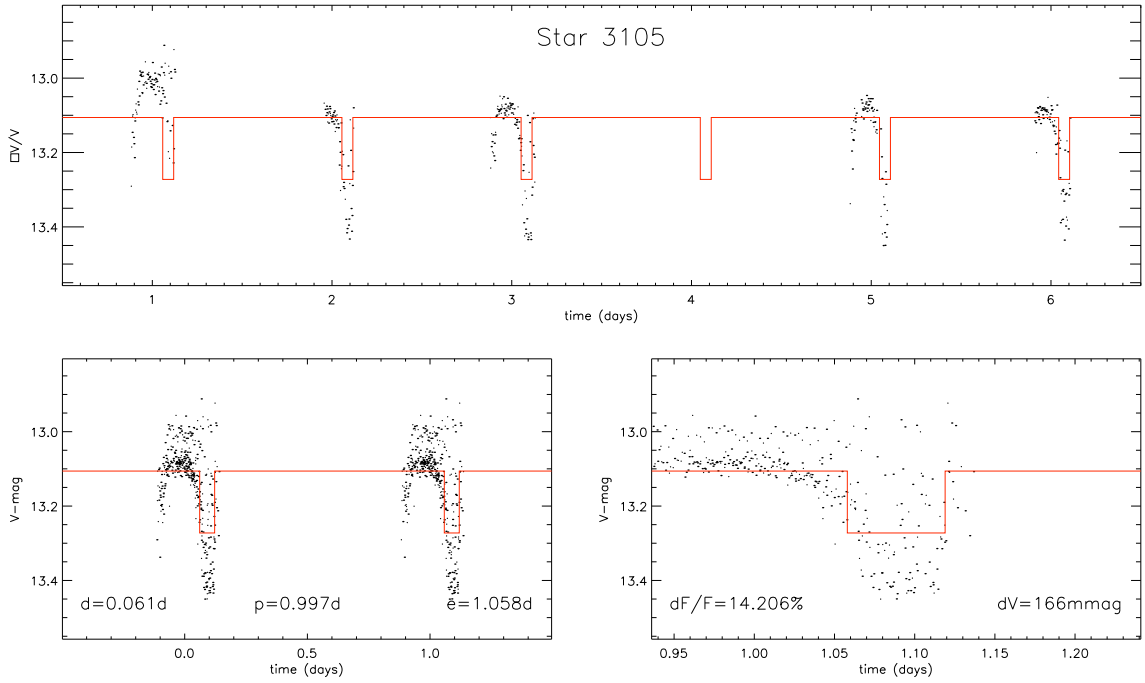

Figure 7.16: Spurious candidate example: night-edge effects.