Draft version August 8, 2019
Typeset using LaTeX **twocolumn** style in AASTeX62

Variability of massive stars in M31 from the Palomar Transient Factory

Monika D. Soraisam,[1] Lars Bildsten,[2,3] Maria R. Drout,[4,5] Thomas A. Prince,[6] Thomas Kupfer,[3] Frank Masci,[7] Russ R. Laher,[7] and Shrinivas R. Kulkarni[6]

[1] *National Optical Astronomy Observatory, Tucson, AZ 85719, USA*
[2] *Department of Physics, University of California, Santa Barbara, CA 93106, USA*
[3] *Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA*
[4] *Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, Ontario, M5S 3H4 Canada*
[5] *The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101,USA*
[6] *Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*
[7] *Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA*

(Received ; Revised ; Accepted)

ABSTRACT

Using data from the (intermediate) Palomar Transient Factory (iPTF), we characterize the time variability of ≈ 500 massive stars in M31. Our sample is those stars which are spectrally typed by Massey and collaborators, including Luminous Blue Variables, Wolf-Rayets, and warm and cool supergiants. We use the high-cadence, long-baseline (≈ 5 years) data from the iPTF survey, coupled with data-processing tools that model complex features in the light curves. We find widespread photometric (R-band) variability in the upper Hertzsprung Russell diagram (or CMD) with an increasing prevalence of variability with later spectral types. Red stars ($V - I > 1.5$) exhibit larger amplitude fluctuations than their bluer counterparts. We extract a characteristic variability timescale, $t_{\rm ch}$, via wavelet transformations that are sensitive to both continuous and localized fluctuations. Cool supergiants are characterized by longer timescales (> 100 days) than the hotter stars. The latter have typical timescales of tens of days but cover a wider range, from our resolution limit of a few days to longer than 100 days timescales. Using a 60-night block of data straddling two nights with a cadence of around 2 minutes, we extracted $t_{\rm ch}$ in the range 0.1–10 days with amplitudes of a few percent for 13 stars. Though there is broad agreement between the observed variability characteristics in the different parts of the upper CMD with theoretical predictions, detailed comparison requires models with a more comprehensive treatment of the various physical processes operating in these stars such as pulsation, subsurface convection, and the effect of binary companions.

Keywords: galaxies: groups: individual (M31) – stars: massive – supergiants – stars: oscillations – catalogs – surveys

1. INTRODUCTION

Massive stars, characterized by initial main-sequence masses greater than around 10 $M_\odot$ and thus being relatively rare and short-lived (Massey 2003, 2013), are important players in varied astrophysical phenomena, shaping the Universe at scales both small and large. For example, their strong stellar winds and radiation fields, and their eventual supernova explosions, cause the chemical enrichment, ionization, and turbulent motion of their local interstellar medium (ISM; Abbott 1982; Freyer et al. 2003, 2006; Matzner 2002; Nomoto et al. 2013); the collective radiation energy of their population acts as a key source of ionization in galaxies (Haffner et al. 2009; Wolfire et al. 2003), their populations in high-redshift galaxies are considered to significantly contribute to the cosmic re-ionization budget (e.g., Robertson et al. 2015), and the collective energy from their final explosions provides vital feedback regulating galaxy formation and evolution (Veilleux et al. 2005; Governato et al. 2010).

Corresponding author: Monika Soraisam
soraisam@noao.edu

These stars are expected to undergo variability at various stages in their evolution—instability in the radiation-dominated envelopes for the hot stars driven by the iron opacity bump (Paxton et al. 2015; Jiang et al. 2015), pulsational instability driven by the opacity of the partial ionization zone of hydrogen for the cool supergiants (Li & Gong 1994; Yoon & Cantiello 2010) as well as the warm supergiants (the classical instability strip; Bono et al. 1999a), and convectively driven low-amplitude stochastic variability both in hot and cool stars (Cantiello et al. 2009; Stothers 2010; Aerts & Rogers 2015; Simón-Díaz et al. 2018). These different instabilities in a star manifest as observable photometric and/or spectroscopic variability. In particular, pulsations play an important role in the evolution of massive stars through their effect on mass loss (e.g., Neilson & Lester 2008; Yoon & Cantiello 2010; Jiang et al. 2018; Ouchi & Maeda 2019). Observations of variability from massive stars, particularly photometric variability, easily afforded by the multitude of current/upcoming time-domain surveys, will provide powerful constraints on their theoretical models. Since it will be impossible to observationally pursue the evolution of a single massive star, statistical studies of large samples of massive stars representing many spectral types are needed to gain important insights into their populations.

A wide range of timescales is expected to characterize the photometric variability of massive stars (e.g., Lovy et al. 1984; Heger et al. 1997; Yoon & Cantiello 2010; Jiang et al. 2015, 2018). Accordingly, wide-field, high-cadence time-domain surveys from the ground with a long baseline extending to a decade, e.g., ASAS (Pojmanski 2002), OGLE (Udalski 2003), and the more recent Palomar Transient Factory (PTF) survey (Law et al. 2009; Rau et al. 2009), are well suited for observational studies of the variability of these stars. In fact, statistical studies of variability have been performed for massive stars in the Large Magellanic Cloud (LMC; Szczygieł et al. 2010) and in the Small Magellanic Cloud (SMC; Kourniotis et al. 2014) based on ASAS and OGLE data, respectively. Space-based surveys (e.g., *HST, Kepler, TESS*), with their exquisite photometry, are desirable, but their narrow field of view and baseline generally limit such analyses to probing short timescales for a handful of stars, with the exception of a few fields (Dorn-Wallenstein et al. 2019; Bowman et al. 2019).

Recently Conroy et al. (2018) presented their work based on *HST* data on the variability of massive stars in M51, which has a high star formation rate. Spetsieri et al. (2018) also used ≈ 3 months of archival *HST* data to catalog massive stars in the Virgo Cluster galaxy NGC 4535 and analyze their variability; Gaia Collaboration et al. (2018b) performed a similar study based on 22 months of *Gaia* data for the Milky Way's stellar populations extending from luminous stars down to dwarf stars. Conroy et al. (2018) extended the baseline to a decade using archival *HST* data of M51. All these studies examined the light curves of the stars to determine the fraction showing observable variability, and extracted variability features such as amplitudes and periods; the latter were also used to label the periodic variables. These quantities were mapped out in the color-magnitude diagram (CMD), and in agreement with our previous study (Soraisam et al. 2018, hereafter Paper I), Conroy et al. (2018) found almost all cool supergiants to be variable. The latter authors also made a direct comparison of the observations with theoretical predictions for instabilities (based on, for example, Paxton et al. 2013, 2015; Yoon & Cantiello 2010) and found agreement for most parts of the CMD, except for some regions occupied by the fainter blue and red variables that require further investigation.

In this paper, we use the wealth of data—both spectroscopic and time-resolved photometric, obtained from Massey et al. (2016), hereafter MNS16, and the PTF survey, respectively—available for the galaxy M31 to study variability of its massive-star population. Given the importance of massive stars in the evolution of galaxies, performing such studies for a large sample of diverse host-galaxy environments (e.g., with varying metallicities and star-formation rates) will inform us both on the physics of the variability phenomena and on their relation to galactic evolution. This work substantially extends the sample size of galaxies with the variability of massive stars characterized in the CMD. Furthermore, timescales of variability were only estimated for periodic light curves in the preceding similar studies. Here, by using a wavelet-based analysis, we also determine the characteristic timescale $t_{\rm ch}$ for variability that may be localized (non-periodic) or unlocalized (periodic) in time.

This paper is organized as follows. In Sect. 2, we describe the sample of massive stars in M31 used for this study and the time-domain data from PTF used for constructing their light curves. We present the method for analyzing the light curves for variabilities and their characteristic timescales in Sect. 3, and show the results of our analysis in Sect. 4. We discuss our results in Sect. 5 and conclude with Sect. 6.

2. DATA

We perform the variability analysis of massive stars in M31 using two complementary data sets: (1) a catalog of spectroscopically typed luminous stars in M31 (MNS16)

and (2) data from the long-baseline optical time-domain survey of M31 by PTF. These data sets are described below.

2.1. *Spectroscopic sample of massive stars in M31*

The Local Group Galaxy Survey (LGGS), by providing an extensive catalog of stars in M31 with multiband photometry (Massey et al. 2006), has enabled drawing up lists of bright and isolated targets for spectroscopic observations and their subsequent characterization (e.g., Massey et al. 2007; Cordiner et al. 2011; Neugent et al. 2012). MNS16 doubled the number of massive stars with spectroscopic labels, particularly those extending from O- to G-type. MNS16 further updated the LGGS catalog of M31, including the spectral classifications from their study as well as those available from earlier efforts; this updated catalog also includes the majority of the spectral classifications for the stars studied by Humphreys et al. (2017). Massey & Evans (2016) also performed a study exclusively focused on the Red Supergiant (RSG) content of M31, contemporaneous with MNS16, obtaining the spectral types for $\approx$ 200 RSGs. This RSG catalog was the source for our study of RSG variability in Paper I. The spectral type information for these RSGs is also included in the updated LGGS catalog.

In total, 1050 stars in the LGGS catalog of MNS16 have spectral classifications with membership flag (Mm) greater than 0, where a value of 0 indicates a foreground star, 1 a confirmed member of M31, 2 a probable member, and 3 unknown membership. These 1050 stars form our reference sample for study of massive star variability in M31, around 70% of which have an Mm flag value of 1. We expect the foreground contamination in the remaining 30% to be minimal; as may be expected, we find many of them to match objects in *Gaia*-DR2 (Gaia Collaboration et al. 2018a, 2016) but the signal-to-noise ratio of the measured proper motions for almost all of the matches is low, indicating that the stars are likely extragalactic. A detailed analysis of foreground contamination is beyond the scope of this paper.

Our reference sample includes (candidate) luminous blue variables (LBVs), Wolf-Rayet stars, OB stars with luminosity classes I, III, and V (for O-type), and supergiants with spectral types ranging from A to M (see MNS16 for more details).

2.2. *M31 PTF time-domain data*

The (intermediate) PTF survey (Masci et al. 2017) provides a rich optical time-domain data set for M31 (Soraisam et al. 2017), with almost daily sampling for about five years. We use the same iPTF/PTF[1] data set in the R-band as in Paper I, covering around 1.8×2.4 deg^2 at depth reaching $m_R \approx 21$, to construct the forced-photometry-based light curves of our 1050 reference sample stars. More details about the survey data as well as the photometry process are given in Paper I and references therein. Once the Zwicky Transient Facility (Bellm et al. 2019), the successor to iPTF, completes operation, the combined baseline will extend to over a decade.

Of the 1050 stars in the reference sample, 1015 are in the PTF footprint that we have analyzed. We construct their light curves via forced photometry on the difference images (see Masci et al. 2017 for details of the iPTF difference-image pipeline). To convert these measured flux differences to absolute fluxes, we add the corresponding subtracted fluxes of the stars obtained by cross-matching (using a search radius of 2″–the typical PTF FWHM) our reference sample with the PSF-fit photometry catalog of the template image used in the subtraction (Laher et al. 2014). For 63 of the stars in our reference sample, we do not find counterparts in the template image catalog within the search radius threshold of 2″. Around 40 of these are tagged as Wolf-Rayets in MNS16. The LGGS R-band magnitudes for almost half of these 63 stars are in the range 18–20 while the remaining half have values greater than 20 mag, which is the typical depth of the iPTF survey. For these 63 stars, the distance to the nearest neighbor in the template image extends up to 8.5″, where the chance of random association increases. It is likely that a combination of faint magnitudes and/or offsets of more than the standard 2″ between the position of the detection on the difference images and the tabulated source position in the template image catalog prevents identification of their counterparts. We drop these 63 sources (just around 5% of the reference sample) from further consideration.

The iPTF light curves for the stars in our reference sample are available at DataLab hosted by the National Optical Astronomy Observatory.[2]

3. ANALYSIS OF LIGHT CURVES

3.1. *Photometric variability*

We use the iPTF R-band light curves of the massive stars constructed above in order to asses their photometric variability. Similar to Paper I, we compute the

[1] We use 'PTF' and 'iPTF' interchangeably in this paper.

[2] The light curve data can be accessed programmatically on DataLab (https://datalab.noao.edu) at soraisam://public/m31PTFmassivestars/

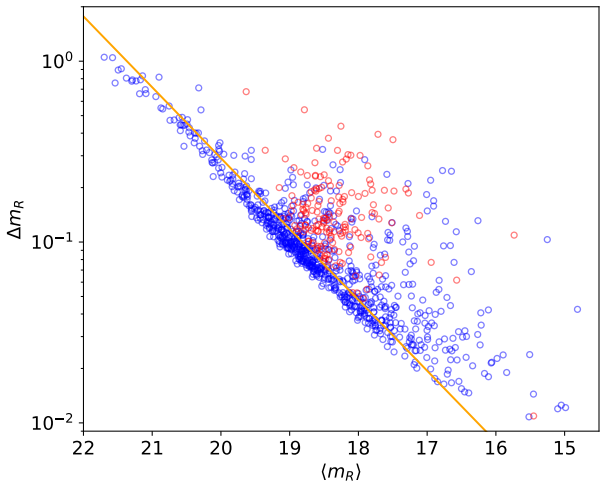

Figure 1. Distribution of standard deviation Δm_R against mean magnitude $\langle m_R \rangle$ for the massive stars in M31 based on iPTF light curves. The red points mark the RSGs with observed variability from Paper I. Points above the orange line are stars with their computed RMS variability exceeding that from noise.

mean magnitude $\langle m_R \rangle$ and standard deviation about the mean (Δm_R) for each star. Figure 1 shows the variation of Δm_R with the mean magnitude of the star. The figure shows absolute variability in magnitude units, equivalent to relative variability in flux units. The orange line shows the noise level of PTF from Paper I, which was computed considering only stars with $\langle m_R \rangle < 20$ and extrapolated to fainter stars. As may be expected, the sensitivity of the PTF survey to relative photometric variability increases with the mean brightness of the stars.

The Δm_R values of the few sources at $\langle m_R \rangle > 20$ above the orange line (Fig. 1) are suspect since the typical depth of the iPTF single-exposure observations is around 20 mag. Indeed, visually examining the light curves of these sources reveals that they are dominated by large error bars. We therefore ignore sources whose mean magnitudes are fainter than 20 mag. Given the sensitivity of the iPTF survey, we find 502 stars with $\langle m_R \rangle < 20$ above the orange line, i.e., with observable variability. Under the spectral type column of MNS16, 8 of these 502 stars are labelled as HII and 3 listed as belonging to clusters without a proper spectral specification. One of the stars belonging to HII regions is likely a W UMa-type contact binary (see Appendix A). We drop these 11 stars, thus bringing our final sample of massive stars with observed variability to 491. For 16 of the stars in our sample, we find the angular separation amongst pairs of them to be less than $5''$, which is comparable to the typical iPTF FWHM of $2''$, and therefore effects of blending are present for them. Nevertheless, we still include them in our analysis, since in this paper we are particularly interested in the population statistics. The total number of stars in our reference sample with $\langle m_R \rangle < 20$ not belonging to clusters or HII regions and with no discernible variability from iPTF is 384.

3.2. *Characteristic timescales for variable stars*

The nature of the photometric variability exhibited by these massive stars is quite diverse: it can be completely stochastic, perfectly periodic, or anything in between. Some examples highlighting the different flavors of variation in the light curves are shown in Appendix B. As mentioned before, a large range in the characteristic timescales of variability of massive stars is expected.

We ascribe the timescale(s) found from a time-frequency or, more precisely, a time-scale length analysis of the light curve using a wavelet transform, as the characteristic timescale(s) of variability for the star, $t_{\rm ch}$. The conventional form of timescale determination based on Fourier analysis (as used in previous studies, see Sect. 1), is limited in its application—it is meant for stationary signals without a time-varying frequency. For periodic signals, Fourier-based analysis is well-justified, but not for a more generic analysis of timescales covering different forms of variation in the light curves. On the other hand, a wavelet transform using oscillatory functions with finite duration, i.e., wavelets, offers a reasonably convenient tool that allows investigating not just a regular mode of variation (i.e., periodicity) but also fluctuations that may be evolving or even transient, along with their timescales via the scale parameter values used to stretch or dilate the wavelet (see Eq. 2).

Despite the efficacy of the wavelet transform formalism, it finds limited application in the analysis of time-series that are unevenly sampled, which is almost always the case for astronomical time-domain data, including iPTF. Some examples of tackling the gaps in data for the purposes of employing a wavelet transform are provided by Foster (1996) and Frick et al. (1997). We take an alternative approach here, by first reconstructing the light curve to fill the gaps and then simply performing a continuous wavelet transform using the Morlet wavelet. We make use of a Gaussian Process model to reconstruct the signal via the critical filter tool of Oppermann et al. (2013) in the NIFTy package of Selig et al. (2013). We successfully implemented this approach in Paper I for analyzing the RSG light curves characterized by varied forms of fluctuations (see Paper I for more details). Some examples of such reconstructions are shown

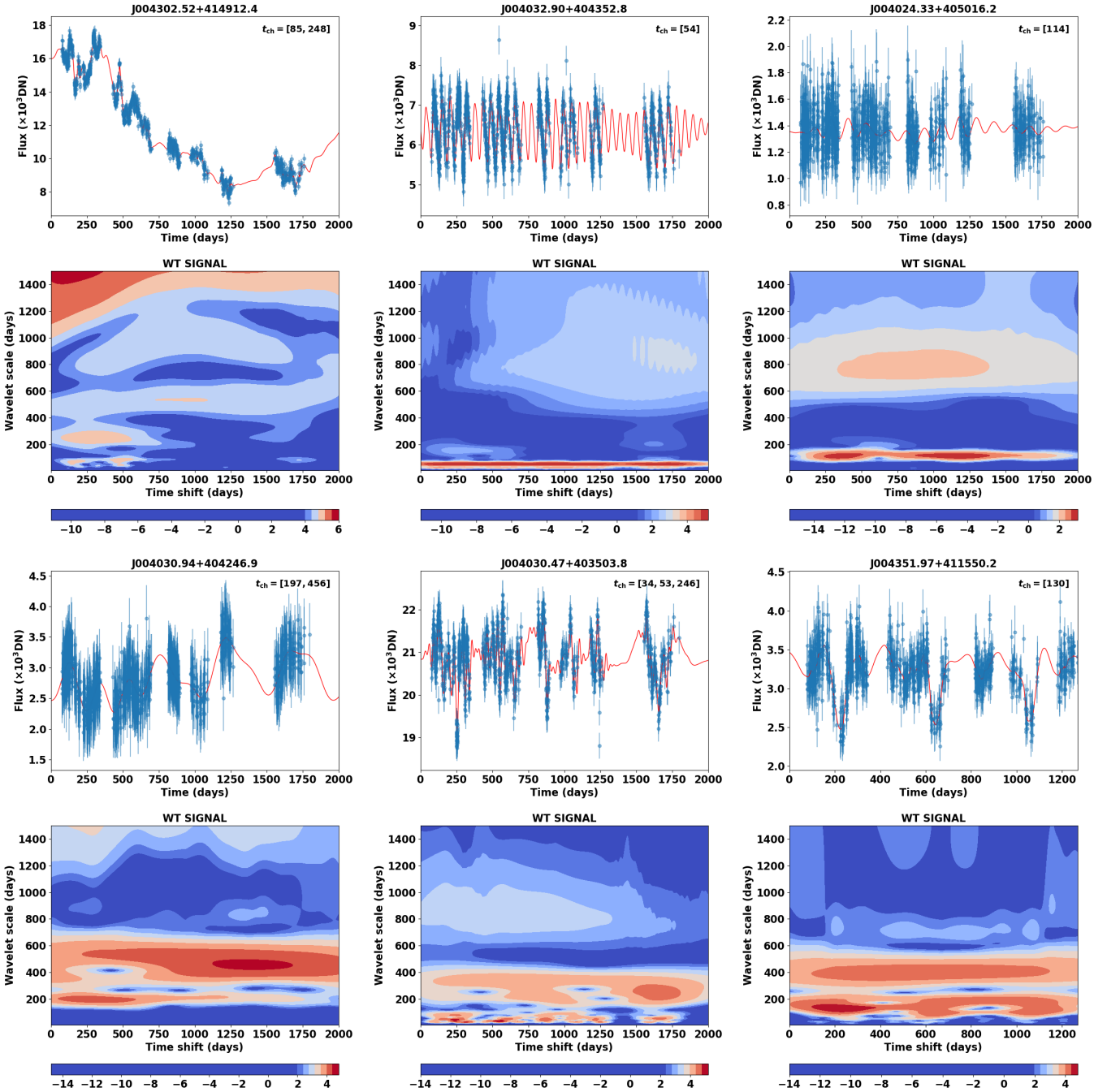

Figure 2. Examples of iPTF light curves along with the corresponding wavelet transform maps for an LBV (cross-id AE And), YSG, and WR star (*top*, left to right), and for supergiants of type O, B, A (*bottom*, left to right). The red curve in each panel with the observed light curve shows the reconstruction (see text). The ID of the star from MNS16 is shown on top of each plot. The time axis in the light curve plots is with respect to a reference value of MJD 56000. The wavelet transform values are shown in log scale.

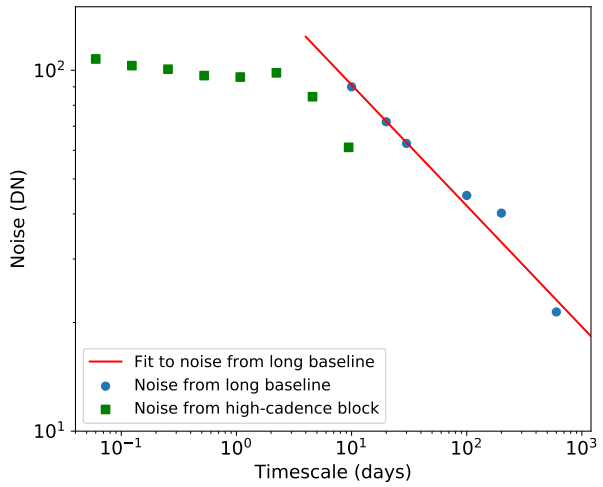

Figure 3. Approximate values of noise (in flux units) for PTF data against characteristic timescales over which they are applicable, obtained by smoothing out the contributions from smaller timescales. The blue circles denote noise estimates extracted using the long baseline PTF light curves of static stars from Paper I and the red line is the fit to them. The green square symbols denote noise estimates obtained using the high-cadence block of the light curves (see Sect. 4.1).

in Fig. 2 and also in Appendix B, which also show the diverse shapes of the light curves of massive stars.

We convolve the reconstructed light curve with the Morlet wavelet, which is a harmonic function with a Gaussian envelope suited for analysis of variable-star light curves, taking scales ranging from 1 day to 1500 days, and extracting the transformation coefficient at every convolution step. The continuous wavelet transform of a light curve $f(t)$ is

$$w(a,b) = a^{-1} \int_{-\infty}^{\infty} f(t')\, \psi^* \left(\frac{t'-b}{a} \right) dt', \qquad (1)$$

$$\psi(t) = e^{\frac{-t^2}{2}} e^{i\omega_o t}, \qquad (2)$$

where a and b represent the wavelet scale and translation (time shift) parameter, respectively, * denotes complex conjugation, and ψ is the Morlet wavelet where we take the dimensionless constant $\omega_o = 6$ such that the wavelet decays significantly in a single cycle (see, for example, Foster 1996; Frick et al. 1997). Note that the transformation is normalized here with $1/a$, instead of the conventional $1/\sqrt{a}$, following Lilly (2017) to ensure that the transform coefficient scales proportionally to the RMS deviation about the mean of the signal, hereafter referred to as amplitude.

The resulting map of the correlation power (transformation coefficient-squared, i.e., $|w(a,b)|^2$) is shown in the bottom panel of the corresponding light curve in Fig. 2, where a large correlation value (indicated in red in the figure) at a given wavelet scale a indicates the presence of such a timescale in the light curve. Because of the uncertainty principle, we cannot both localize the frequency and time. Therefore, significant power at small wavelet scales (with better time resolution) appears as narrow streaks or blobs; in the alternative case, power excess at large wavelet scales corresponding to poor time resolution, result in big blobs in the maps. These features are evident in the bottom-middle panel of Fig. 2.

We extract the characteristic timescales $t_{\rm ch}$ of the stars from the correlation-power maps, by identifying connected regions of power above a threshold, which we define as 5σ above the background fluctuations (the background is obtained by taking the mean of the pixel values lower than the 75th percentile of the distribution and σ is their standard deviation). For each of the resulting connected regions, we identify the wavelet scale outside the cone of influence[3] corresponding to the pixel with the maximum power. We term this scale a as $t_{\rm ch}$. If a star's wavelet transform map has no region above its corresponding threshold, then no timescales will be reported for it, whereas if there are multiple islands of excess power above the background, then we will obtain multiple timescales. Because of this automated detection of $t_{\rm ch}$, in some cases, an apparently contiguous region of excess wavelet transform power could be broken up into multiple regions.

We further filter the extracted values of $t_{\rm ch}$ on the basis of their corresponding *amplitudes*. We define the amplitude for a characteristic timescale $t_{\rm ch}$ of the given light curve, hereafter termed as $t_{\rm ch}$-specific amplitude, as $\alpha\, P^{1/2}$, where P is the wavelet-transform correlation power evaluated at its local maximum. The constant α is chosen such that if the signal is itself a wavelet with timescale $t_{\rm ch}$, this amplitude matches the RMS of the signal within $\pm t_{\rm ch}$ of its center (see Appendix C). The $t_{\rm ch}$-specific amplitudes are then compared with the expected noise at similar timescales. The timescale-specific noise values are estimated by using the light curves of the static/non-varying stars from Paper I and smoothing out the noise contributions from smaller timescales. At a given timescale t we smooth the light curves by binning them using bin widths of size $t/2$ and compute the RMS deviations (Δflux) of the smoothed light curves, assuming uncorrelated noise. Finally, we

[3] Cone of influence is the region in the wavelet transform map affected by edge effects given that we are dealing with finite time series.

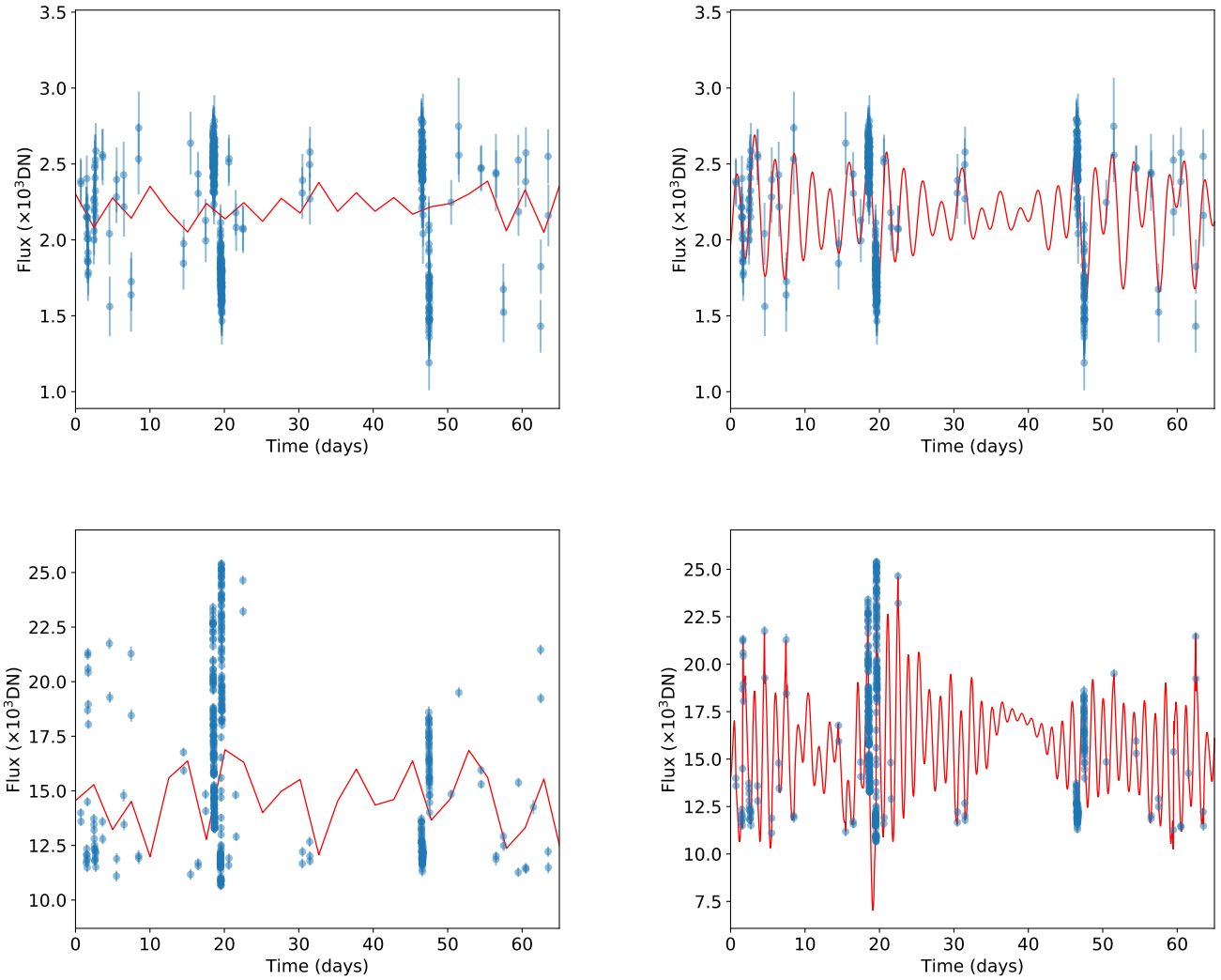

Figure 4. iPTF light curves for the two stars—J004026.84+403504.6 (*top*) and J004509.86+413031.5 (*bottom*)—vetted out as discussed in Sect. 3.2. Their wavelet transforms are shown in Fig. 15 of Appendix B. The reconstructed signals, shown in red, in their respective left panels are generated with a maximum resolution of 3 days, while the ones in the right panels are for a resolution of around 30 minutes constructed using the high-cadence block (see text). The time axes in all panels are shown with respect to a reference MJD of 56250.621783.

take the median of the Δflux values of the static stars to define the noise at timescale t. These noise estimates are shown as a function of the timescales they are applicable to in Fig. 3. We compare the amplitude obtained for a $t_{\rm ch}$ to the corresponding noise and drop those $t_{\rm ch}$ values with associated variability consistent with noise (i.e., amplitude below the red line in Fig. 3).

The maximum resolution of the reconstructed light curve is ≈ 3 days for the long baseline that we are considering. Increasing the resolution is not possible because of computational limitations. Hence, we only consider timescales recovered from the maps that are larger than 10 days and therefore well sampled at our resolution. To ascertain that the reconstructions represent the data reasonably well, we visually examine all reconstructed light curves for which $t_{\rm ch}$ was obtained. In two cases, the star consistently shows large amplitude fluctuations on a timescale less than our maximum resolution of 3 days. One of these two stars, J004509.86+413031.5, is labeled as a Yellow Supergiant candidate (YSG:) in the MNS16 catalog and the other, J004026.84+403504.6, as a composite spectral type O7+O9f: without a luminosity class (':' denotes candidate in MNS16). Their light curves are shown in Fig. 4, where the data taken on a few nights with an ultra-high cadence of around 2 minutes highlight the insufficient modelling resolution. For a uniform analysis of all stars, we drop the two sources J004026.84+403504.6 and J004509.86+413031.5

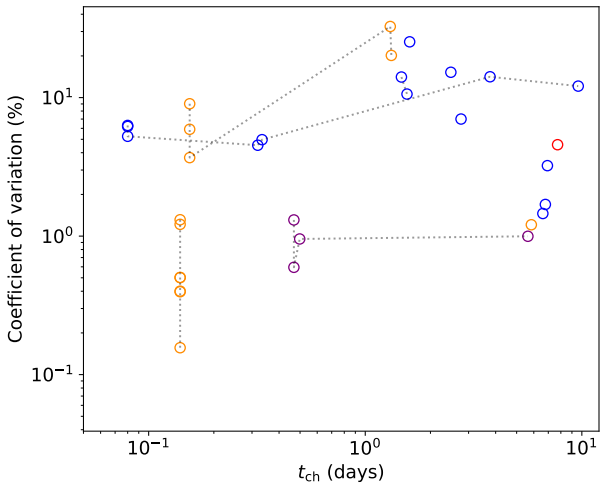

Figure 5. Coefficient of variation, i.e., the ratio of the localized RMS amplitude to the average flux value of the light curve, as a function of the extracted t_{ch} for the 13 stars having significant variations in the high-cadence blocks. The data points from the same star are connected by a dotted line. The O- and B-type stars are shown in blue, A-type supergiant in purple, yellow supergiants in orange, and M-type supergiant in red.

when evaluating for $t_{ch} \geq 10$ days, thereby decreasing our sample size to 489 (however, see the following section for determining $t_{ch} < 10$ days). Of these, we determine $t_{ch} \geq 10$ days for 356 stars. Typically there are two timescales from a single star for those with any timescales. A tally of the stars is given in Table 1.

3.3. *High-cadence data*

In light of the few high-cadence nights included in the PTF data set, we repeated the t_{ch} analysis considering only around 60 nights as the baseline, straddling all the high-cadence ones, and increasing the maximum resolution in the light curve reconstruction to ≈ 30 minutes. We term the data from this shortened baseline as the high-cadence block.

For the majority of the stars (around 63% of the 491 sources), the fluctuations on this short baseline are hidden under the noise (orange line in Fig. 1). For those above the noise floor (182 stars), we perform wavelet transformation of their reconstructed light curves from the high-cadence block. This set of 182 stars includes J004026.84+403504.6 and J004509.86+413031.5 discussed above, and we obtain reasonable reconstructions using their high-cadence blocks, as shown in the right panels of Fig. 4.

We extract t_{ch} from the wavelet-correlation-power maps obtained using the high-cadence block signal, in a similar manner as in Sect. 3.2. As can be seen from Fig. 3, the noise estimates for the short t_{ch} values from the high-cadence block are approximately constant (≈ 100 DN); the noise estimate is biased for the long smoothing timescale similar to the baseline considered (60 days), and hence there is a sharp drop around those timescales (t_{ch} around 10 days). We use a constant threshold of 100 DN to filter short t_{ch} values with corresponding amplitudes consistent with noise. The results are described in the next section.

4. RESULTS

4.1. *Short characteristic timescales from the high-cadence block*

We find significant variations with $t_{ch} < 10$ days for 13 stars, two of which are J004026.84+403504.6 and J004509.86+413031.5. The reconstructed light curves along with the wavelet-correlation-power maps for the 13 stars are shown in Fig. 15 of Appendix B.

For J004026.84+403504.6, we obtain t_{ch} from the high-cadence block of around 2.5 days, while for the YSG candidate J004509.86+413031.5, t_{ch} around 0.15 days and 1.3 days (see Fig. 15). As can be seen, the light curve of this candidate YSG is quite regular and its t_{ch} values are consistent with periods of Cepheids, hence it is likely a Cepheid in M31 and its variability can be attributed to pulsation.

Dorn-Wallenstein et al. (2019) recently studied the variability of seven massive stars, comprising YSGs and LBVs, in the LMC based on *TESS* data, which continuously monitored the stars with a 2-minute cadence for ≈ 27 days. They find short-timescale (< 10 days) variability, from analysis of periodograms (hence periodic timescales), for five of the stars with amplitudes $< 1\%$ albeit without constraining the nature of variability. In agreement with their result, we also find existence of short-timescale variability in evolved stars in M31. For the larger sample size of stars (182) and their data with two times longer baseline we have analyzed, we find significant short t_{ch} values (either periodic and/or non-periodic) for $\approx 7\%$ of the stars in our sample. The spectral types of the remaining 11 stars consist of supergiants of O (including a main-sequence), A, B, M-types and yellow supergiants (Fig. 15).

In Fig. 5, we show the amplitudes as a function of t_{ch} for the 13 stars. For these stars, the variability amplitude at t_{ch} values $\lesssim 1$ day is $\lesssim 10\%$, while for t_{ch} between 1 and 10 days, the amplitude ranges from 1% to a few $\times 10\%$. Given that the ground-based PTF data have much larger noise than the *TESS* data, we can observe low levels of variability ($< 1\%$) only for the brightest stars. On the other hand, we find high levels of short-

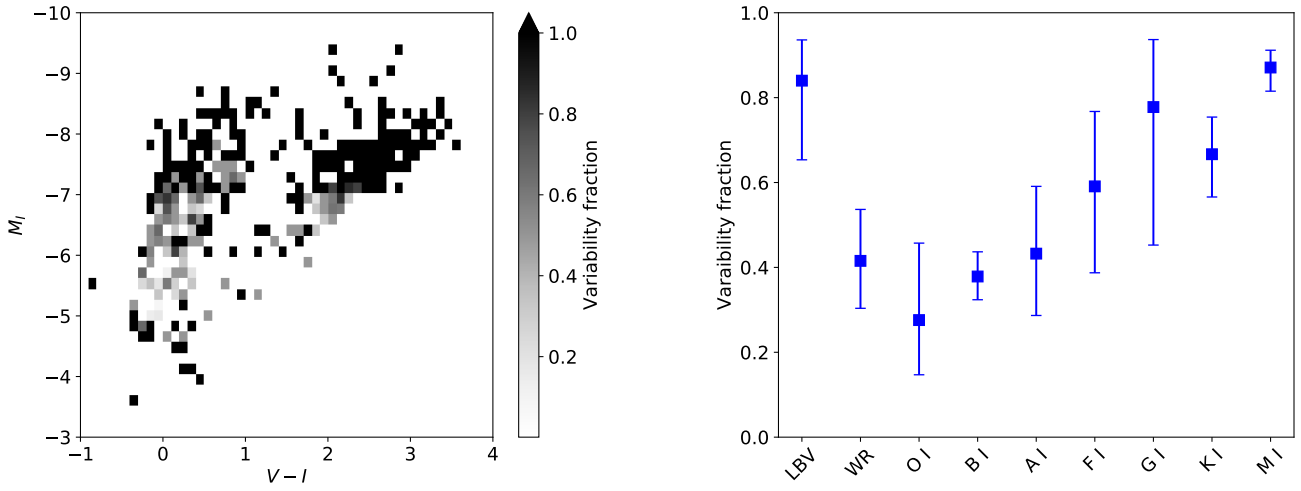

Figure 6. CMD (*left*) showing the fraction of massive stars in M31 identified by MNS16 with observed variability on timescales ≥ 10 days from iPTF, and as a function of spectral types for the supergiants (*right*). The vertical error bars in the right panel indicate the 95% binomial confidence interval.

Table 1. Summary of our reference sample of massive stars in M31

Total spectroscopic sample	1050
In the PTF footprint	1015
Brighter than $m_R = 20$ and with PTF-detectable variability	502
Long-timescale wavelet analysis performed	489
Significant $t_{\rm ch} \geq 10$ days detected	356
Short-timescale wavelet analysis performed	182
Significant $t_{\rm ch} < 10$ days detected	13

timescale variability ($> 1\%$–10%) only for $< 7\%$ of the stars in our much larger sample. For the small number of stars Dorn-Wallenstein et al. (2019) analyzed, it is thus not surprising that they did not find such variability.

Higher-cadence data (and thus with reduced noise) over an extended baseline for a large sample of stars will be ideal for probing short(er) timescales in these stars, even for the fainter ones.

4.2. *Results based on the long baseline light curve data*

The variability characteristics we have determined for the massive stars in M31 are listed in Table 2. We also list in this table all $t_{\rm ch} > 10$ days extracted automatically for a given star, with the caveat mentioned in Sect. 3.2 that some of these $t_{\rm ch}$ values reported for a star, in particular, when the values are close to one another, may result from segmentation of a single connected region in our automated analysis.

Using the available multiband photometry from LGGS for the massive stars in M31 studied here, coupled with the variability information we have extracted from their time-domain iPTF data (Table 2), we now map out their variability characteristics in the CMD. In particular, we consider the B, V, and I photometric measurements from LGGS. However, it is to be noted that the LGGS photometries do not represent long-baseline averaged values. Nevertheless, these LGGS-CMD-based variability maps still provide information for the population as a whole because any variable photometry will average out when looking at the whole population.

4.2.1. *Map of variability fraction in the upper CMD of M31*

We show the variability fraction in the $V-I$ vs. M_I CMD in the left panel of Fig. 6, which can be directly compared with Fig. 13 of Conroy et al. (2018), and also as a function of the spectral types of supergiants extending from O- to M-type, and including LBVs and Wolf-Rayets, in the right panel of the same figure. In computing both the colors and absolute magnitudes, we have corrected for the foreground extinction using the reddening along the line of sight to M31, $E(B-V) = 0.062$ from Schlegel et al. (1998); we do not account for the interstellar extinction intrinsic to M31. We use a dis-

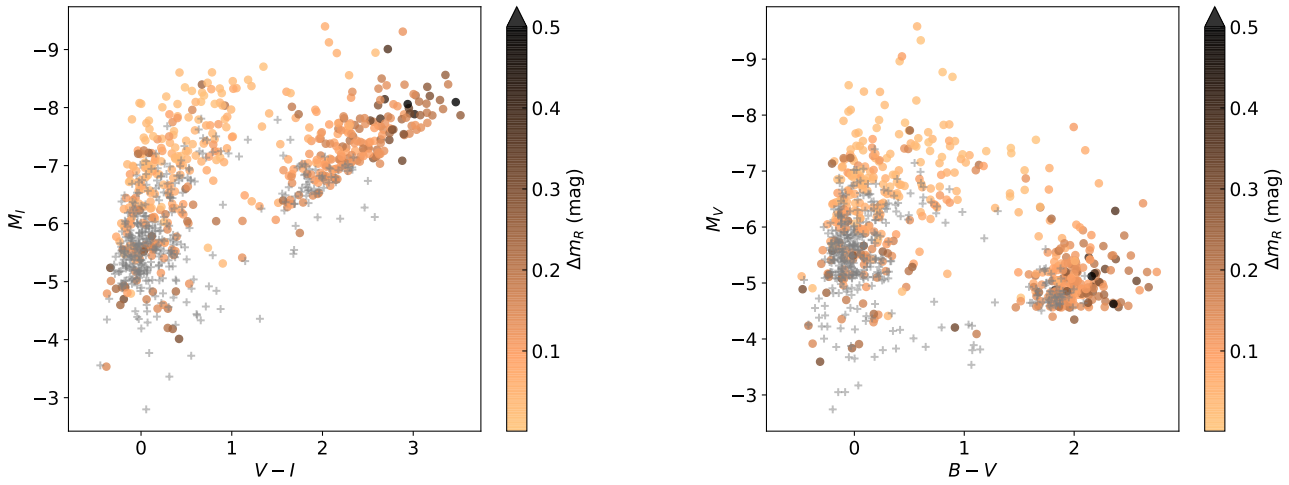

Figure 7. CMDs based on the LGGS photometry of the massive stars in M31 from MNS16, color-coded by the observed variability amplitude from iPTF, expressed as the RMS deviation Δm_R from the mean of the light curve. Grey plus symbols signify objects that did not have detectable variability in our data.

tance modulus of 24.36 for M31 (Vilardell et al. 2010). Note that the selection effects of our parent data sets propagate into our study—the spectroscopic catalog of MNS16 is likely complete only for the Wolf-Rayets, and the sensitivity of the time-domain PTF data to variability drops for the fainter stars to > 0.1 mag (Fig. 1). This caveat applies to all results presented in the following.

Conroy et al. (2018) covered a much larger parameter space in the CMD, particularly the evolved states of late-type main-sequence stars including AGBs. These are not represented in our study since the MNS16 work (and the companion work on RSGs; Massey & Evans 2016) was primarily directed toward massive stars in regions where contamination from the foreground is minimal. This is also the reason for the gap, where there are no stars in our reference sample, between the two groups in the left panel of Fig. 6. Further, the sample of stars used by Conroy et al. (2018) is complete for the luminous stars, but they probed variability at a level > 0.03 mag comparable to our study. The very high variability fractions we have found for the brightest stars are not reliant on the increased sensitivity of PTF to photometric variability for such bright stars, since these variations tend to have large amplitudes (see Sect. 4.2.2, Fig. 7).

For the parts of the CMDs analyzed here, there is good agreement between our results and those of Conroy et al. (2018), including the unexplained variability observed in faint blue stars (around $0 < V-I < 0.5$ and $M_I > -8$). Those stars are located between the classical instability strip and the region comprising instability in radiation-dominated envelopes of blue massive stars around the location of iron-opacity peaks (Jiang et al. 2015) deduced by Conroy et al. (2018) based on a simple parameterization of the instability in 1D evolutionary models. The increase in the variability fraction of these stars toward later types can be clearly seen in the right panel of Fig. 6, reaching almost 100% for the M supergiants, in agreement with Conroy et al. (2018). LBVs by definition are variables (see Sect. 4.2.3), hence, their fraction is close to 100%. We caution, however, that the result in Fig. 6 is subject to incompleteness as described above, particularly for the early type O-stars. But it is also to be noted that variability properties were not part of the selection function for the MNS16 spectral catalog.

4.2.2. *Map of variability amplitude for massive stars in M31*

Figure 7 shows the aforementioned variability maps in two CMDs (M_I vs. $V-I$ and M_V vs. $B-V$), where the amplitude of fluctuation is expressed as the RMS deviation from the mean of the light curve Δm_R. As can be seen, the stars without observed variability from iPTF (those below the orange line in Fig. 1; the grey points in the present plot) are largely concentrated at the faint part of the CMD, in particular the blue stars. This is in line with their low-SNR data, and hence we obtain only upper limits on their variability.

Two groups are clearly evident in this figure–among other things, they reflect the selection cut adopted by MNS16 for their spectroscopic follow-up. The group redward of around 1.0 for either color $V-I$ or $B-V$ mainly corresponds to RSGs, while the other group consists of warm (spectral type A, F, G) supergiants and blue stars that include both evolved and non-evolved ones. Our

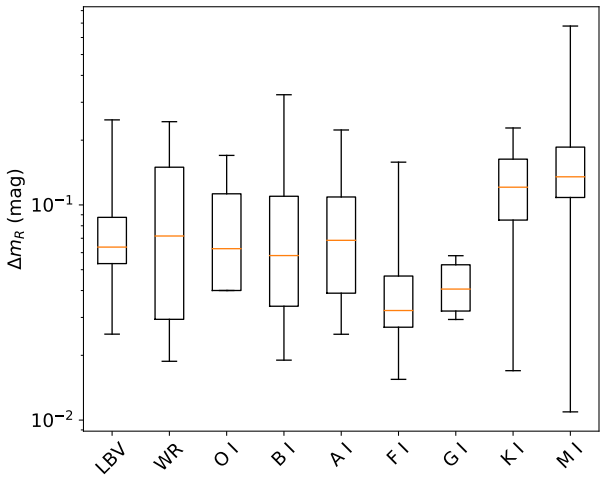

Figure 8. Range of RMS amplitudes for the different spectral types of supergiants in M31. The box extends from the first quartile to the third quartile of the corresponding RMS distribution, and the orange line shows the median while the whiskers extend to the minimum and maximum values of the distribution.

variability maps thus show that the red evolved stars typically exhibit larger variability amplitude than the bluer counterparts on timescales ≥ 10 days.

Figure 8 shows the RMS amplitude distributions as a function of the spectral types of supergiants. Note that all the different timescales of variability of the star contribute to this amplitude. It is clear that the later stellar types, specifically K and M supergiants, exhibit larger amplitudes of photometric variation than the early-type supergiants. This result also agrees with that of Conroy et al. (2018) for M51. Theoretical predictions of brightness amplitudes for massive stars that can be directly compared with our results are not available in the literature. On the other hand, our observational results will help to constrain models making such computations in the future.

4.2.3. *Photometric variability of M31 LBVs (and candidates)*

LBVs are evolved hot stars located in the upper HR diagram with luminosity close to the Eddington limit, generally characterized by the prototype S Dor-like variability and sometimes giant eruptions like that of η Car accompanied by enhanced mass-loss (Hubble & Sandage 1953; Humphreys & Davidson 1994). They represent a rare class of variable stars, with less than around 50 of them identified in the Local Group, and their evolution, including the driving mechanism of variability, are not well understood.

Our analysis also allows us to study the photometric variability of the LBVs and candidates in the MNS16 compilation. There are 25 such LBVs in total in their Table 3. Four of them are below the detection threshold of variability for our study (i.e., the orange line in Fig. 1). For the remaining 21 with observed variability, we show all light curves in Fig. 14 in Appendix B. This includes 17 candidate LBVs. Thus, this study establishes photometric variability for the majority of the LBV candidates for the first time.

The RMS amplitudes of the LBVs extracted from the nearly five-year-long PTF data are at most a few tenths of a magnitude (cf. Fig. 8). This is in contrast to the S Dor- and η Car-type variabilities in LBVs that have amplitudes greater than around 1 mag and timescales on the order of years to decades – the amplitudes and timescales being greater for the latter type of eruptions (e.g., Humphreys & Davidson 1994). However, LBVs also exhibit a third kind of variability, namely stochastic microvariations with amplitudes on order of 0.1 mag in timescales of days to even hundreds of days (e.g., Abolmasov 2011; Mehner et al. 2017).

Recently, Jiang et al. (2018) performed 3D hydrodynamical simulations of radiation-dominated envelopes for LBVs and found that convection (owing to the opacity peaks in the envelopes) drives irregular variations in these stars with characteristic timescales of a few days for (10–30)% variability in luminosity; their simulations covered only less than 1000 hours of the envelope evolution and the variability level includes the contributions from smaller timescales as well. One of the theoretical light curves computed by Jiang et al. is shown in Fig. 16. Smoothing out the contributions of timescales $\lesssim 10$ days, the variability is at the 5% level for the longer timescales.

The extracted characteristic timescales, $t_{\rm ch}$ for these stars in M31 (Sect. 4.2.4, Fig. 10) cover a wide range from tens of days to a thousand days, with the distribution concentrated toward $t_{\rm ch}$ values of a few tens of days. The $t_{\rm ch}$-specific amplitudes for this class of stars are around a few percent as shown in Fig. 11, in agreement with the theoretical results of Jiang et al. (2018). This points to the observed variability in these LBVs associated with the extracted $t_{\rm ch}$ values as likely corresponding to stochastic microvariability.

4.2.4. *Map of $t_{\rm ch}$ for massive stars in M31*

Similar to the maps we have constructed for the amplitude of variability, we also map out the characteristic timescale $t_{\rm ch}$ in Fig. 9. As mentioned in Sect. 3.2, we obtain multiple $t_{\rm ch}$ values for many of the stars. In this figure, we show only one timescale per star correspond-

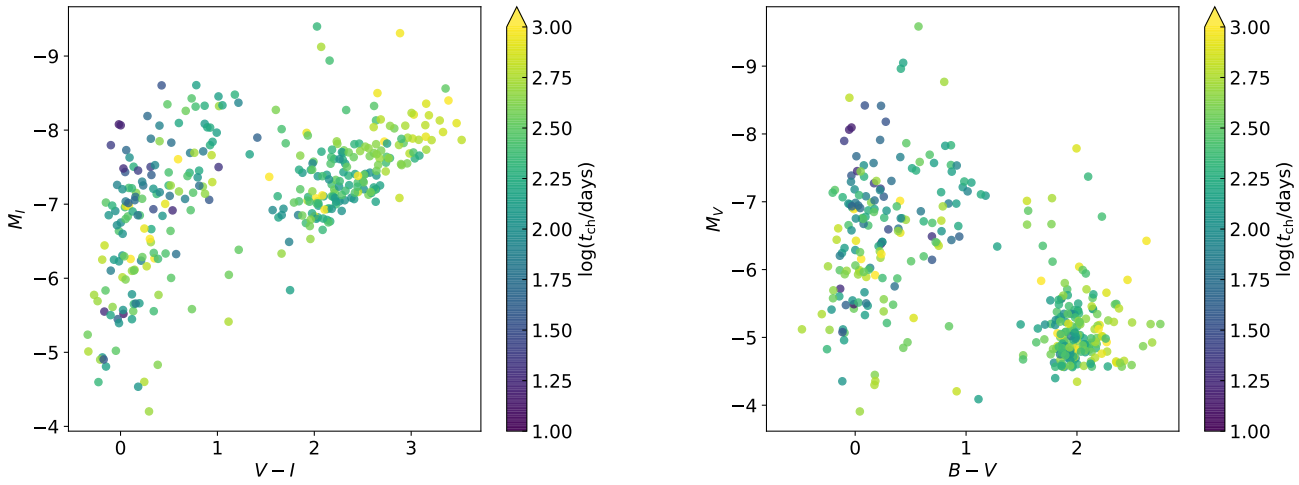

Figure 9. CMDs similar to Fig. 7, where the color-coding in this case reflects a single characteristic timescale for each star corresponding to the global maximum power in the star's wavelet transform.

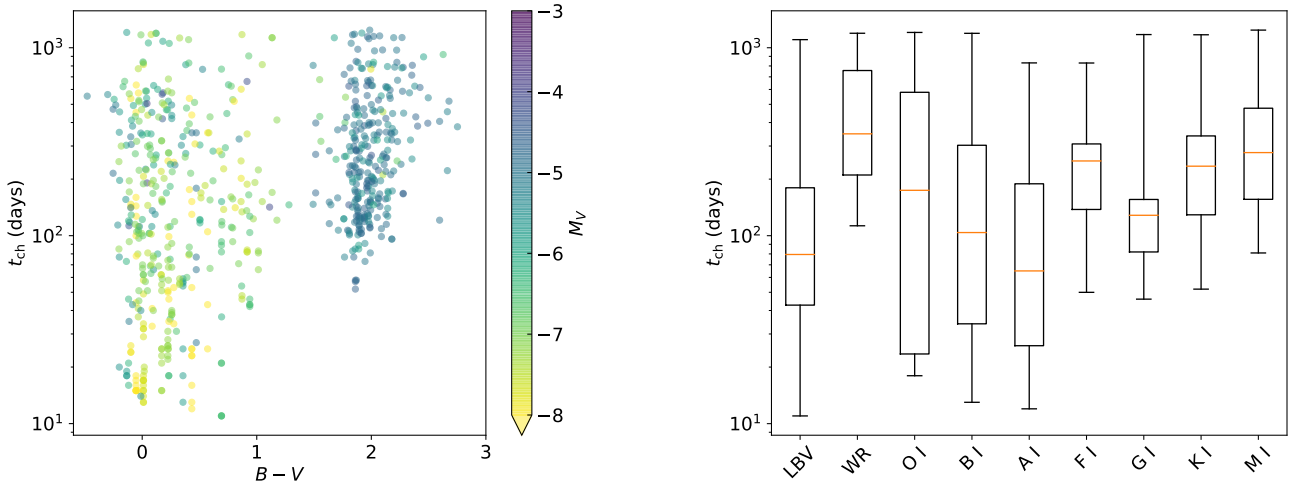

Figure 10. *Left*: Characteristic timescales t_{ch}, obtained from the wavelet transforms of the iPTF light curves (Sect. 3.2), against color obtained from the LGGS photometry of massive stars in M31. It is to be noted that the density of points should not be over-interpreted, as our automated determination of connected regions with high wavelet transform values could in some cases fragment a connected region. *Right*: Range of t_{ch} values for different types of supergiants in M31. The box extends from the first quartile to the third quartile values of the corresponding distribution, and the orange line shows the median while the whiskers extend to the minimum and maximum values. The upper limit of t_{ch} probed here is $\approx$ 1200 days, imposed by the baseline of the iPTF survey in our wavelet transform analysis, while there is a lower limit of t_{ch} = 10 days dictated by the maximum resolution of our light curve reconstruction.

ing to the maximum power in the wavelet transform of its light curve. As is evident from the plot, the group of red massive stars is characterized by longer t_{ch} than the bluer stars; the latter group shows a comparatively large range of timescales. This result is highlighted even further in Fig. 10 (left panel), where we show all extracted timescales against the $B-V$ color, and in the right panel of the same figure, where we show the t_{ch} distribution for the different types of supergiants in M31 with observed variability. The lower (around 10 days) and upper limits (around 1200 days) in t_{ch} in these figures are due to constraints set by the maximum resolution of the Gaussian Process modeling of the light curve (Sect. 3.2) and the baseline of the survey, respectively.

It can be seen from the right-hand panel of Fig. 10 that the LBVs and the A-type supergiants have typical $t_{\rm ch}$ values of a few tens of days, shorter than the yellow (F- and G-types) and red supergiants (K- and M-types), which typically have $t_{\rm ch}$ values of a few hundred days. We find that the WRs, O-, and B-type supergiants in our sample have typical $t_{\rm ch} \gtrsim 100$ days; however, we caution that the sample incompleteness for these spectral types could affect our results for them. For almost all the types, the distribution of $t_{\rm ch}$ values extends to the longest timescale (> 1000 days) probed in our study, while, as discussed above, the distributions for the bluer stars, i.e., LBVs, O-, B-, and A-types, also extend (close) to the shortest timescale probed here.

For the cool massive stars, long pulsation timescales on the order of a few hundred to a few thousand days are predicted (e.g., Heger et al. 1997), and this appears consistent with our observed $t_{\rm ch}$ values for these stars. In Paper I, we extracted periodic timescales from spectroscopically-confirmed RSGs and derived their period-luminosity relation in M31. We found multiple periodic timescales for many of these stars as evidenced by the multiple peaks in their power spectra ranging from a few hundred to a thousand days—the maximum timescale that could be probed in that study (see Paper I). The earlier results are thus also in agreement with those from the current study (cf. Fig. 10).

Furthermore, the observed trend in the timescales for the various types of supergiants (right panel of Fig. 10) broadly agrees with the simple predictions made by Lovy et al. (1984). Theoretical results from recent, more sophisticated, 3D hydrodynamical simulations of LBVs by Jiang et al. (2018) are also in-line with the observations, where the $t_{\rm ch}$ values of the LBVs are concentrated at short timescale values (right panel of Fig. 10; Sect. 4.2.3). It will be interesting to compare future 3D simulations covering a much longer timescale with our observational results.

In Fig. 11, we show the amplitude associated with a given $t_{\rm ch}$ relative to the average flux of the star (i.e., coefficient of variation) for the different types of supergiants. In almost all cases, the trend between the coefficient of variation and $t_{\rm ch}$ appears flat. However, note that since the noise decreases with increase in timescale probed (cf. Fig. 3), we are likely more incomplete at the shorter $t_{\rm ch}$ values than at the longer values. Furthermore, a very low level of variability, especially at short $t_{\rm ch}$ values, will be contributed by the brightest stars.

It can be seen from the figure that the level of variability for the K- and M-type supergiants corresponding to the timescales found for them ($t_{\rm ch} \gtrsim 100$ days) is typically a few percent to tens of percent, and these stars exhibit the highest level of variability amongst all supergiants. For the yellow supergiants, the $t_{\rm ch}$-specific variability is $\approx 0.3\%$ to a few percent, while for the O, B-, and A-types, it ranges between $\approx 0.3\%$ to 10%. For WRs, the coefficient of variation corresponding to their $t_{\rm ch} \gtrsim 100$ days is between 0.1% to a few percent. For LBVs, the variability level is generally around a few percent.

5. DISCUSSION

The variability characteristics, comprising amplitudes and timescales, for the massive stars in M31 are empirically derived in this work. What is/are then the physical processes responsible for the variability observed in these stars?

Various processes are at work in massive stars. These include rotation, pulsation, convection, binary interaction, variable stellar winds and the consequent mass loss that are common for these massive stars, and possibly all of them influencing each other.

The binary fraction for RSGs is not yet well-established. Patrick et al. (2019) estimated a binary fraction of 30% for these stars – Neugent et al. (2019) found 63 RSG+B star binaries in M31 and M33 out of a sample size of 149 consistent with the fraction above (one of their stars J004327.01+412808.7 typed as BI in MNS16 is included in our sample and we find a $t_{\rm ch}$ of around 150 days for this star; cf. Fig. 13). Because of their large physical size, however, the binary companions of these stars are expected at large orbital periods – thousands of days, which are much longer than their extracted timescales from the PTF data. RSGs also rotate very slowly—the projected rotational velocity mapped for Betelgeuse based on ALMA observations by Kervella et al. (2018) is around 5 km/s and thus the rotation period is on the order of decades, again much longer than the baseline of the data studied here. However, in Paper I, we found that the extracted periodic timescales for these stars are consistent with the fundamental and first overtone modes of radial pulsations, based on theoretical models computed with MESA (Paxton et al. 2015) coupled with the linear asteroseismology code `GYRE` (Townsend & Teitler 2013). The $t_{\rm ch}$ values for these red massive stars in the present study are similar to those found in Paper I and hence, are likely associated with pulsation as well.

Like for the RSGs, the binary fraction of the rarer yellow supergiants (F- and G-types) is also highly uncertain. For a few cases, orbital periods of around a few hundred days have been determined (Prieto et al. 2008; Sperauskas et al. 2014). Besides, when (some of) these stars go through the classic instability strip, they will

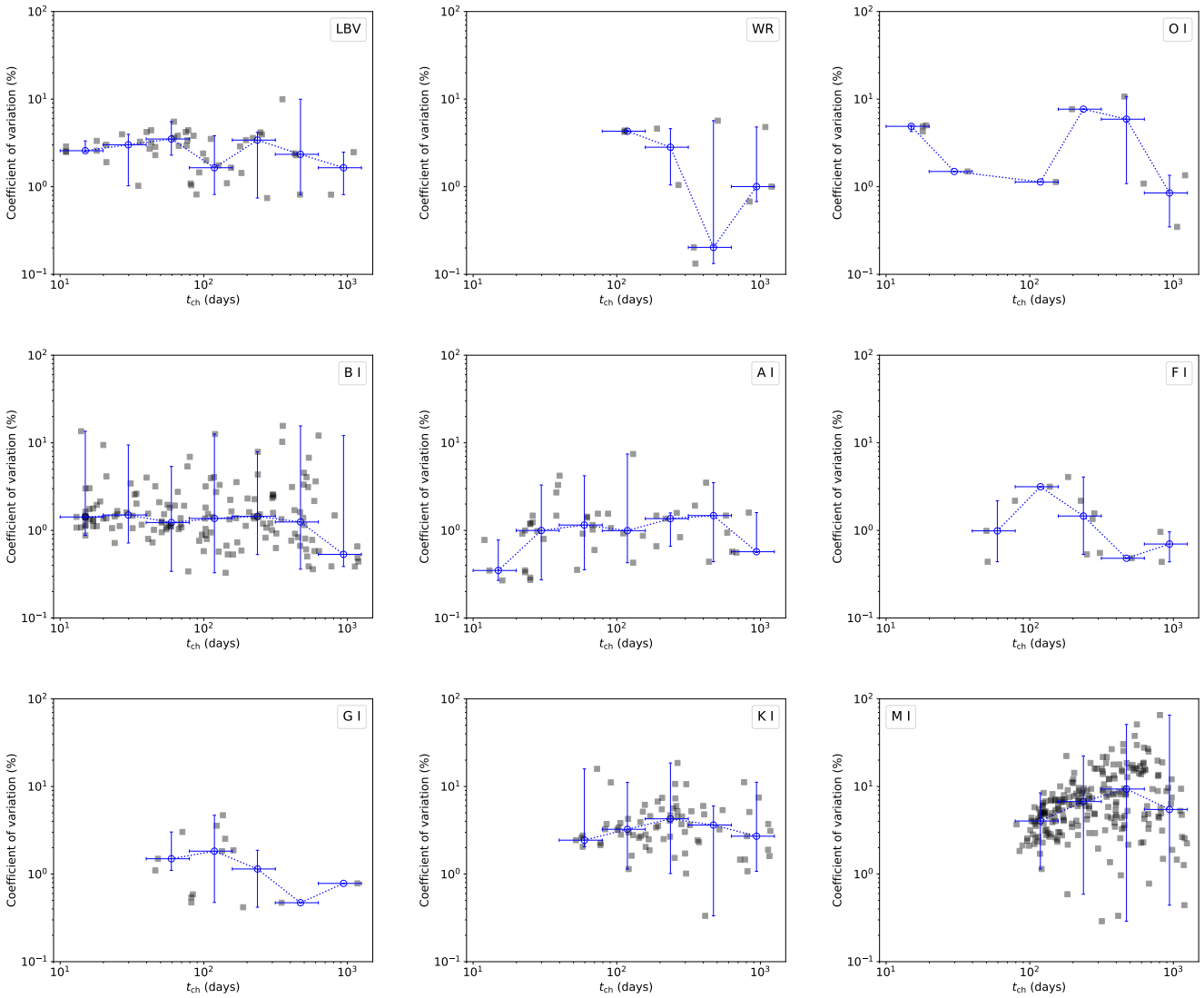

Figure 11. Coefficient of variation as a function of characteristic timescales t_{ch} for the different types of supergiants as indicated in the respective legends. The grey squares are the data points, while the blue circles denote the median after binning the data logarithmically into 7 bins between t_{ch} values of 10 days and 1250 days, using the same bins for each panel. The vertical whiskers extend to the minimum and maximum values of the coefficient of variation at a given t_{ch} bin and the horizontal bar shows the bin width.

undergo pulsation as Cepheids and their variants such as the double-mode Cepheids (Bono et al. 1999b) with periods on the order of a few days to tens of days, and even reaching over 100 days for the ultra-long period Cepheids (Bird et al. 2009). The observed variability in these stars (see e.g., Fig. 13) likely includes pulsation, but disentangling various other processes, like the presence of a binary companion, is impossible without complementary information from multi-epoch long-baseline spectroscopy and comprehensive theoretical models.

For hot luminous stars, the binary fraction is better determined, which is around 50%–70% (e.g., Sana et al. 2012; Dunstall et al. 2015). The period distribution for these stars indicate that orbital periods of tens to hundreds of days, i.e., the timescales covered in our study, are well populated, with the O-types preferring short periods (Sana et al. 2012) and the distribution of B-types being flat (Dunstall et al. 2015). Aerts et al. (2018) performed a study of the blue supergiant ρ Leo using *K2* photometric data and multi-epoch HERMES spectroscopic data, and found dominant variability at the level of 8 mmag with a periodic timescale of a few tens of days, which they attributed to rotation. The shorter t_{ch} values in Fig. 11 and their corresponding amplitudes for the O and B stars may indeed be connected to rotation. Coherent pulsations triggered by the κ effect of

the Fe opacity bump are known to occur in these stars (e.g., Cox et al. 1992; Dziembowski et al. 1993). Given the high metallicity of M31, opacity-driven pulsations are certainly operating in these hot stars. Thus, some, if not all, of the variability we have observed for these stars may be coherent pulsations.

Pedersen et al. (2019) performed a classification of the photometric variability in a large sample (over 150) of O- and B-type stars, which included rotating, eclipsing, and pulsating stars, based on *TESS* data and visually examining the light curves and their discrete Fourier transforms. Many of the variable stars in the Pedersen et al. sample show simultaneous modulations from the different phenomena, e.g., eclipsing light curves with rotational modulation and/or coherent pulsation. Further, a number of blue supergiants in their sample show stochastic low-frequency variability, similar to that found by Bowman et al. (2019) in their analysis of more than 150 hot luminous stars in the ecliptic and the LMC using *K2* and *TESS* data. The characteristic amplitudes of this variability are less than a few millimags (see Bowman et al. 2019). Such variability could be caused by internal gravity waves excited by turbulent core convection (Bowman et al. 2019) or sub-surface convection triggered by local opacity enhancements associated with Fe and He (e.g., Cantiello et al. 2009; Cantiello & Braithwaite 2019). Being a low-frequency phenomenon, it should manifest at long timescales. Thus, the long $t_{\rm ch}$ values of around hundreds to a thousand days that we find for these stars in M31 and the corresponding amplitude of $< 1\%$ (Fig. 11) may be related to such stochastic variability.

A consistent treatment and prediction of observables accounting for the various physical processes, e.g, interaction between pulsation and convection, mass loss, binary effects, is currently lacking to interpret the wealth of observational results. Studies like ours, however, will provide important constraints in modeling the poorly understood physical processes in the evolution of massive stars.

6. CONCLUSIONS

By mining the well-sampled, long-baseline iPTF time-domain data of M31, we have mapped the variability of stars in the upper part of the HR diagram. The earlier work of MNS16 and also Massey & Evans (2016) in M31 provided the identification of the massive stars, including their spectral types. These stars exhibit a wide variety of light curve shapes, encoding the varied physical phenomena modulating their observed radiation fields. In agreement with Conroy et al. (2018), who studied the variability of the stellar populations in M51, we find that (photometric) variability is widespread in the upper parts of the CMD of M31 with the observed variability fraction increasing for the later spectral types toward close to 100% for the cool supergiants. The incompleteness of the spectral catalog, however, likely affects the result for the early-type supergiants. In the observed variability maps, the cooler stars also show larger variability amplitudes than the bluer counterparts.

Further, using the powerful signal reconstruction tool of Oppermann et al. (2013), we are able to extract characteristic timescales $t_{\rm ch}$ of variability for these stars that are both localized in time (e.g., in the case of irregular and semi-regular variables) and unlocalized (i.e., periodicity). For the first time, we are thus able to map out the $t_{\rm ch}$ values for the massive star population characterized by diverse variability behavior.

Using a block of the time-series data straddling two nights with a high cadence of ≈ 2 minutes, we find significant $t_{\rm ch}$ in the range 0.1–10 days for 13 stars in our sample. This is in agreement with recent results from space-based data, for example from Dorn-Wallenstein et al. (2019) that also found such short timescales in a smaller sample of evolved stars.

Using the long baseline PTF light curve for probing $t_{\rm ch} \gtrsim 10$ days, we find that the cool supergiants have longer $t_{\rm ch}$ (hundreds of days and more) relative to the hotter stars. The $t_{\rm ch}$ values of the latter cover a larger range, exhibiting variability typically on short timescales of tens of days and extending to the larger timescale domain of the cooler stars. These observations are in general agreement with the theoretical predictions of pulsation in massive stars, especially for the cooler supergiants. For the hot luminous stars, a myriad of effects including rotational modulation, pulsation, binary companions, and perhaps an interplay among them can result in photometric and/or spectroscopic variability. Contemporaneous multi-epoch spectra will greatly complement studies of variability in massive stars. On the theoretical side, various uncertainties in the treatment of physical processes operating in the massive stars' envelopes (interplay of convection and pulsation, stellar wind, companion effect, etc.) in the models prevent a more detailed comparison with observations.

The maps of the variability characteristics such as those presented in this paper will serve as a powerful tool to explore the phenomena themselves, but also in investigating their relation to the host galaxy environment and consequently galaxy evolution, given the important role that these massive stars play. To this end, a large statistical sample of the maps of variability characteristics of stellar populations in different host environments covering a large range of metallicity, star formation rate,

etc., will be invaluable. Fortunately, this is a task that is now possible to accomplish with the large amount of archival time-domain data that have become available from the present generation of wide-field, high-cadence, and long-baseline optical surveys.

Acknowledgments: This research was supported in part by the National Science Foundation through grant PHY-1748958 at the KITP and benefited from interactions that were funded by the Gordon and Betty Moore Foundation through Grant GBMF5076. MS thanks Chien-Hsiu Lee for helpful discussions on observational studies of stellar variability as well as Niels Oppermann, on proper applications of signal analysis tools. We also thank Charlie Conroy and Yan-Fei Jiang for discussions, and the latter for providing his simulation data. MRD acknowledges support from the Dunlap Institute at the University of Toronto and the Canadian Institute for Advanced Research (CIFAR). This work was based on observations obtained with the 48-inch Samuel Oschin Telescope at the Palomar Observatory as part of the Palomar Transient Factory project. Operations were conducted by Caltech Optical Observatories and data processing by IPAC.

Software: `numpy` (van der Walt et al. 2011), `scipy` (Jones et al. 2001–), `astropy` (Astropy Collaboration et al. 2013), `pyfits`, `pandas` (McKinney 2010), `mpi4py` (Dalcín et al. 2005), `matplotlib` (Hunter 2007), `DAOPHOT` (Stetson 1987), `DAOGROW` (Stetson 1990), `NIFTy` (Selig et al. 2013).

REFERENCES

Abbott, D. C. 1982, ApJ, 263, 723

Abolmasov, P. 2011, New Astronomy, 16, 421

Aerts, C., & Rogers, T. M. 2015, ApJL, 806, L33

Aerts, C., Bowman, D. M., Símon-Díaz, S., et al. 2018, MNRAS, 476, 1234

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002

Bird, J. C., Stanek, K. Z., & Prieto, J. L. 2009, ApJ, 695, 874

Bono, G., Caputo, F., Castellani, V., & Marconi, M. 1999a, ApJ, 512, 711

Bono, G., Marconi, M., & Stellingwerf, R. F. 1999b, ApJS, 122, 167

Bowman, D. M., Burssens, S., Pedersen, M. G., et al. 2019, Nature Astronomy, 320

Cantiello, M., & Braithwaite, J. 2019, arXiv e-prints, arXiv:1904.02161

Cantiello, M., Langer, N., Brott, I., et al. 2009, A&A, 499, 279

Chen, X., Deng, L., de Grijs, R., Wang, S., & Feng, Y. 2018, ApJ, 859, 140

Conroy, C., Strader, J., van Dokkum, P., et al. 2018, ApJ, 864, 111

Cordiner, M. A., Cox, N. L. J., Evans, C. J., et al. 2011, ApJ, 726, 39

Cox, A. N., Morgan, S. M., Rogers, F. J., & Iglesias, C. A. 1992, ApJ, 393, 272

Dalcín, L., Paz, R., & Storti, M. 2005, Journal of Parallel and Distributed Computing, 65, 1108

Dorn-Wallenstein, T. Z., Levesque, E. M., & Davenport, J. R. A. 2019, arXiv e-prints, arXiv:1901.09930

Dunstall, P. R., Dufton, P. L., Sana, H., et al. 2015, A&A, 580, A93

Dziembowski, W. A., Moskalik, P., & Pamyatnykh, A. A. 1993, MNRAS, 265, 588

Foster, G. 1996, AJ, 112, 1709

Freyer, T., Hensler, G., & Yorke, H. W. 2003, ApJ, 594, 888

—. 2006, ApJ, 638, 262

Frick, P., Baliunas, S. L., Galyagin, D., Sokoloff, D., & Soon, W. 1997, ApJ, 483, 426

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018a, A&A, 616, A1

Gaia Collaboration, Eyer, L., Rimoldini, L., et al. 2018b, arXiv e-prints, arXiv:1804.09382

Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203

Haffner, L. M., Dettmar, R.-J., Beckman, J. E., et al. 2009, Reviews of Modern Physics, 81, 969

Heger, A., Jeannin, L., Langer, N., & Baraffe, I. 1997, A&A, 327, 224

Hubble, E., & Sandage, A. 1953, ApJ, 118, 353

Humphreys, R. M., & Davidson, K. 1994, PASP, 106, 1025

Humphreys, R. M., Davidson, K., Hahn, D., Martin, J. C., & Weis, K. 2017, ApJ, 844, 40

Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90

Jiang, Y.-F., Cantiello, M., Bildsten, L., Quataert, E., & Blaes, O. 2015, ApJ, 813, 74

Table 2. Variability characteristics of massive stars in M31

ID[a]	SpT[b]	$\langle m_R \rangle$	Δm_R	$t_{\rm ch}$ (days)[c]
J004314.06+415301.8	B8I	17.75	0.05	223, 407
J004353.34+414638.9	WN7	18.28	0.12	-99
J004043.95+405901.6	G8I	17.34	0.06	71
J004028.48+404440.2	B9I	17.23	0.03	37, 41, 112, 773
J004157.56+410753.3	A4I	17.68	0.05	62, 189, 830
J004213.75+412524.7	M3 I	18.57	0.13	144, 362, 846
J004313.31+413329.1	F0I	18.48	0.10	274, 812
J004252.10+414516.4	K5 I	19.02	0.13	235, 781
J004621.08+421308.2	cLBV	17.76	0.05	36, 75, 183
J004331.17+411203.5	WC8	19.08	0.17	-99
J004130.30+411603.8	A7I	17.29	0.11	25, 58, 68, 276, 682
J004009.43+405932.3	ON9.7Iab	18.93	0.17	18, 19, 1206
J003953.55+402827.7	YSG	17.28	0.21	170, 1130, 1131
J004158.87+405316.7	O9.5I	17.98	0.06	37, 152

[a] LGGS ID of star as given in MNS16.

[b] Spectral type of the star from MNS16.

[c] Measurements not available are indicated -99.

(This is only a part of the table; the full version is available in machine-readable form online.)

Jiang, Y.-F., Cantiello, M., Bildsten, L., et al. 2018, Nature, 561, 498

Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python, , , http://www.scipy.org/

Kervella, P., Decin, L., Richards, A. M. S., et al. 2018, A&A, 609, A67

Kourniotis, M., Bonanos, A. Z., Soszyński, I., et al. 2014, A&A, 562, A125

Laher, R. R., Surace, J., Grillmair, C. J., et al. 2014, PASP, 126, 674

Law, N. M., Kulkarni, S. R., Dekany, R. G., et al. 2009, PASP, 121, 1395

Li, Y., & Gong, Z. G. 1994, A&A, 289, 449

Lilly, J. M. 2017, arXiv e-prints, arXiv:1703.04882

Lovy, D., Maeder, A., Noels, A., & Gabriel, M. 1984, A&A, 133, 307

Masci, F. J., Laher, R. R., Rebbapragada, U. D., et al. 2017, PASP, 129, 014002

Massey, P. 2003, ARA&A, 41, 15

—. 2013, New Astronomy Reviews, 57, 14

Massey, P., & Evans, K. A. 2016, ApJ, 826, 224

Massey, P., McNeill, R. T., Olsen, K. A. G., et al. 2007, AJ, 134, 2474

Massey, P., Neugent, K. F., & Smart, B. M. 2016, AJ, 152, 62

Massey, P., Olsen, K. A. G., Hodge, P. W., et al. 2006, AJ, 131, 2478

Matzner, C. D. 2002, ApJ, 566, 302

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51 – 56

Mehner, A., Baade, D., Groh, J. H., et al. 2017, A&A, 608, A124

Neilson, H. R., & Lester, J. B. 2008, ApJ, 684, 569

Neugent, K. F., Levesque, E. M., Massey, P., & Morrell, N. I. 2019, ApJ, 875, 124

Neugent, K. F., Massey, P., & Georgy, C. 2012, ApJ, 759, 11

Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, ARA&A, 51, 457

Oppermann, N., Selig, M., Bell, M. R., & Enßlin, T. A. 2013, PhRvE, 87, 032136

Ouchi, R., & Maeda, K. 2019, arXiv e-prints, arXiv:1904.07878

Patrick, L. R., Lennon, D. J., Britavskiy, N., et al. 2019, A&A, 624, A129

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15

Pedersen, M. G., Chowdhury, S., Johnston, C., et al. 2019, ApJL, 872, L9

Pojmanski, G. 2002, AcA, 52, 397

Prieto, J. L., Stanek, K. Z., Kochanek, C. S., et al. 2008, ApJ, 673, L59

Rau, A., Kulkarni, S. R., Law, N. M., et al. 2009, PASP, 121, 1334

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Selig, M., Bell, M. R., Junklewitz, H., et al. 2013, A&A, 554, A26

Simón-Díaz, S., Aerts, C., Urbaneja, M. A., et al. 2018, A&A, 612, A40

Soraisam, M. D., Gilfanov, M., Kupfer, T., et al. 2017, A&A, 599, A48

Soraisam, M. D., Bildsten, L., Drout, M. R., et al. 2018, ApJ, 859, 73

Sperauskas, J., Začs, L., Raudeliūnas, S., Musaev, F., & Puzin, V. 2014, A&A, 570, A3

Spetsieri, Z. T., Bonanos, A. Z., Kourniotis, M., et al. 2018, A&A, 618, A185

Stetson, P. B. 1987, PASP, 99, 191

—. 1990, PASP, 102, 932

Stothers, R. B. 2010, ApJ, 725, 1170

Szczygieł, D. M., Stanek, K. Z., Bonanos, A. Z., et al. 2010, AJ, 140, 14

Townsend, R. H. D., & Teitler, S. A. 2013, MNRAS, 435, 3406

Udalski, A. 2003, AcA, 53, 291

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing In Science & Engineering, 13, 22

Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARA&A, 43, 769

Vilardell, F., Ribas, I., Jordi, C., Fitzpatrick, E. L., & Guinan, E. F. 2010, A&A, 509, A70

Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278

Yoon, S.-C., & Cantiello, M. 2010, ApJL, 717, L62

APPENDIX

A. W URSA MAJORIS-TYPE CONTACT BINARY CANDIDATE

For the 11 stars belonging to clusters and HII regions, and without proper spectral labels, which we have dropped in Sect. 3.1, we also perform both the high- and low-resolution reconstruction of their light curves to determine $t_{\rm ch}$. Interestingly for one of the stars, J004259.31+410629.1, belonging to an HII region and with Mm of 1 in the MNS16 catalog (Sect. 2.1), we obtain $t_{\rm ch}$ of 0.14 day based on its high-cadence block—its low-resolution reconstruction does not represent the data well (see Fig. 12). Very short periods (less than around 0.5 day) with sinusoidal light curve shape are typical characteristics of W UMa-type contact binaries—both features exhibited by J004259.31+410629.1, such that it may be indeed a W UMa-type contact binary.

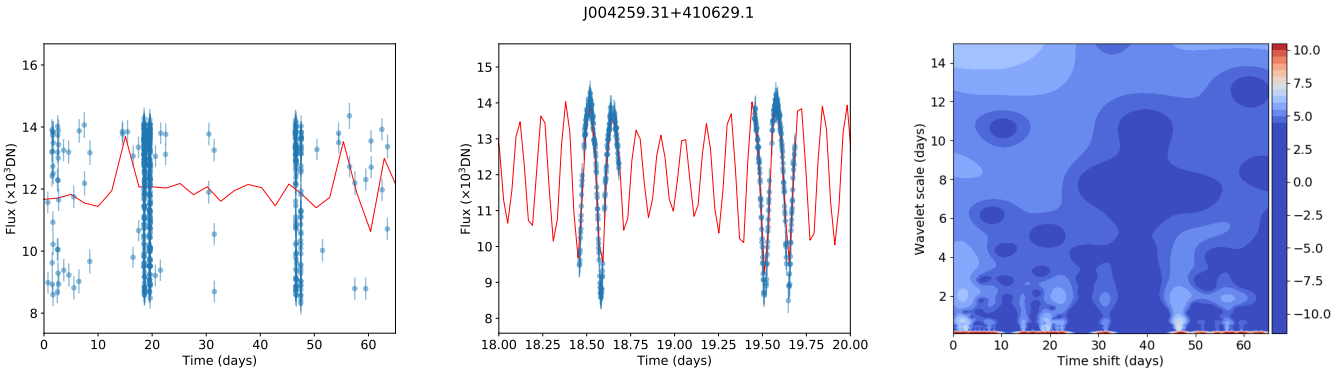

Figure 12. iPTF light curve for the W UMa candidate along with its wavelet transform based on the high-cadence block (*right*). The reconstructed signals are shown in red for both the lower and higher resolution in the left and middle panels, respectively, and the wavelet transform corresponds to the signal in the middle panel.

The mean magnitude of the iPTF light curve of this star is $m_R \approx 17.0$. Applying the optical R band period-luminosity relation for W UMa-type contact binaries from Chen et al. (2018), we find its absolute magnitude $M_R \approx 7.07$, and accordingly a distance estimate of around 0.97 kpc. We find a source in *Gaia*-DR2 within 0.27″ of J004259.31+410629.1, with measured parallax of 0.527 mas. Its estimated distance based on the *Gaia* parallax is around 2 kpc. There is roughly a factor 2 discrepancy between the two distance estimates. Nevertheless, given the number of uncertainties, the discrepancy may not be surprising. It thus appears likely that J004259.31+410629.1 is in the foreground and a contact binary.

B. LIGHT CURVES AND THEIR WAVELET TRANSFORM

Examples of iPTF light curves for stars belonging to different spectral types are shown in Fig. 13, while light curves for all (candidate) LBVs in M31 with detected photometric variability from iPTF are shown in Fig. 14. Light curves of the 13 stars with significant $t_{\rm ch}$ in the high-cadence block are also shown in Fig. 15 along with their corresponding wavelet transform maps.

C. EXAMPLES OF WAVELET TRANSFORM OF SIMULATED SIGNALS

We perform the wavelet transform of some simulated signals with localized events to verify the recovery of the input characteristic timescales $t_{\rm true}$. To this end, we model the localized event with a wavelet and consider three light curves containing one, two and three wavelets, as shown in Fig. 16. Each light curve extends over a baseline of 2000 days with a cadence of 2.5 days. The wavelet signal in the first light curve has an amplitude (in linear flux units) value of 100 and $t_{\rm true}$ of 200 days; for the second light curve, the two wavelets have the same amplitude of 100 but different $t_{\rm true}$ values of 300 days and 30 days; while the third light curve has three wavelets superposed with $t_{\rm true}$ values of 20 days, 100 days, and 50 days and corresponding amplitudes 50, 10, and 100.

The respective wavelet transform power is shown below each simulated light curve, and the recovered timescale $t_{\rm ch}$ applying the same method as in Sect. 3.2 is shown in the legend of each panel. As can be seen, we recover timescales similar to the input values for the three light curves. For the third light curve, the smallest amplitude $t_{\rm ch}$ is missed.

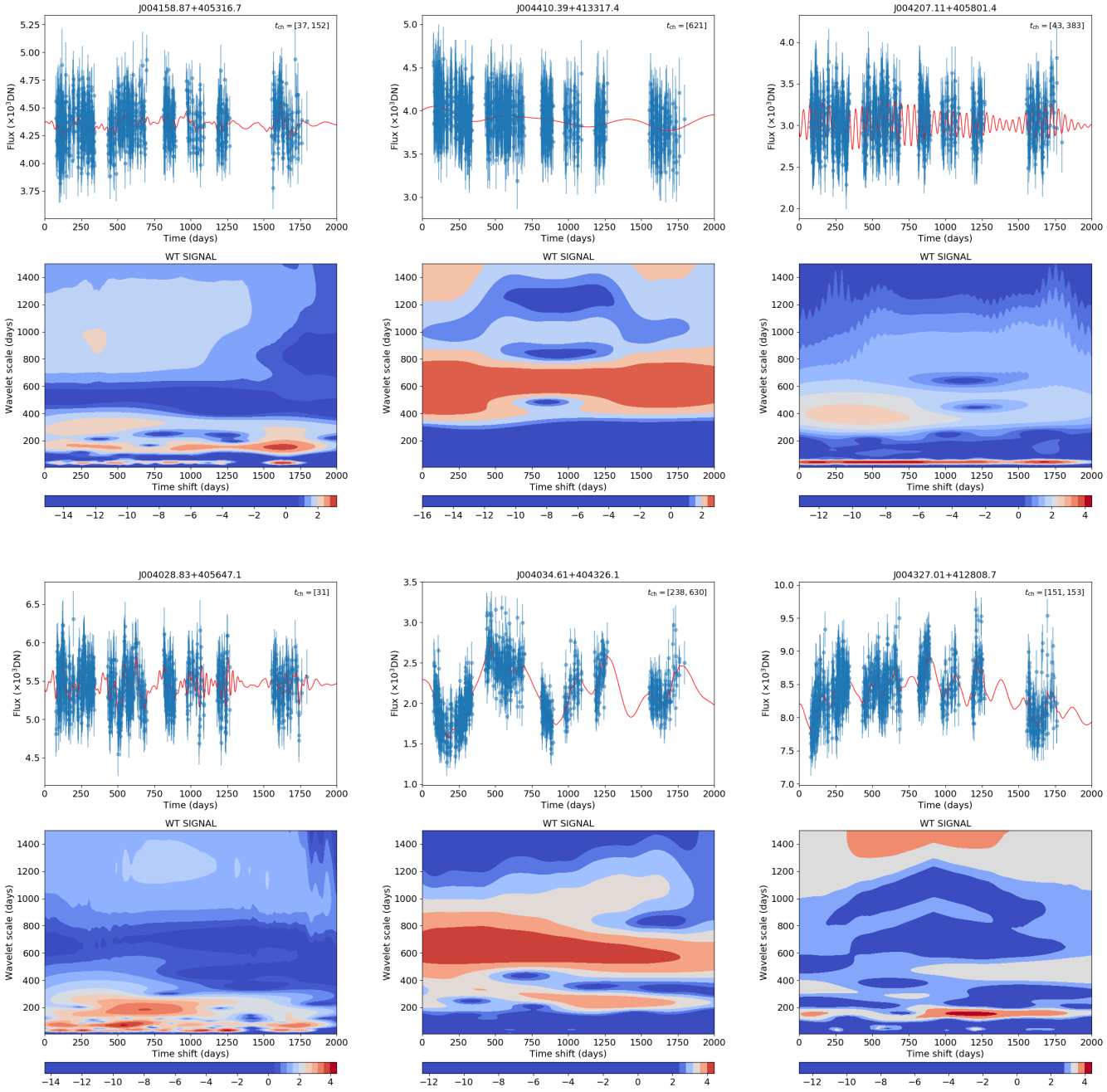

Figure 13. iPTF light curves along with the corresponding wavelet transform maps for O stars (*top*) and B stars (*bottom*). The red curve shows the reconstructed signal, and the ID of the star from MNS16 is indicated on top of the plot.

This is due to the fact that, for the adopted background threshold, the island of power excess for this timescale has merged with the neighboring one. We experimented with different threshold values, but the threshold adopted in Sect. 3.2 appears optimal and is hence used throughout our analysis.

We find that the absolute magnitude of the wavelet transform coefficient, or the square root of the transform power, corresponding to the recovered timescale varies with the amplitude of the signal and inversely with the pixel size used in the Gaussian-Process reconstruction of the light curve. We determine the scaling factor by comparing the RMS amplitude of the simulated light curve, evaluated over an interval of width $2 \times t_{\rm true}$, with the root of the transform power corresponding to $t_{\rm ch}$, and find it to be $1.7 \times$ pixel/2.5(days).

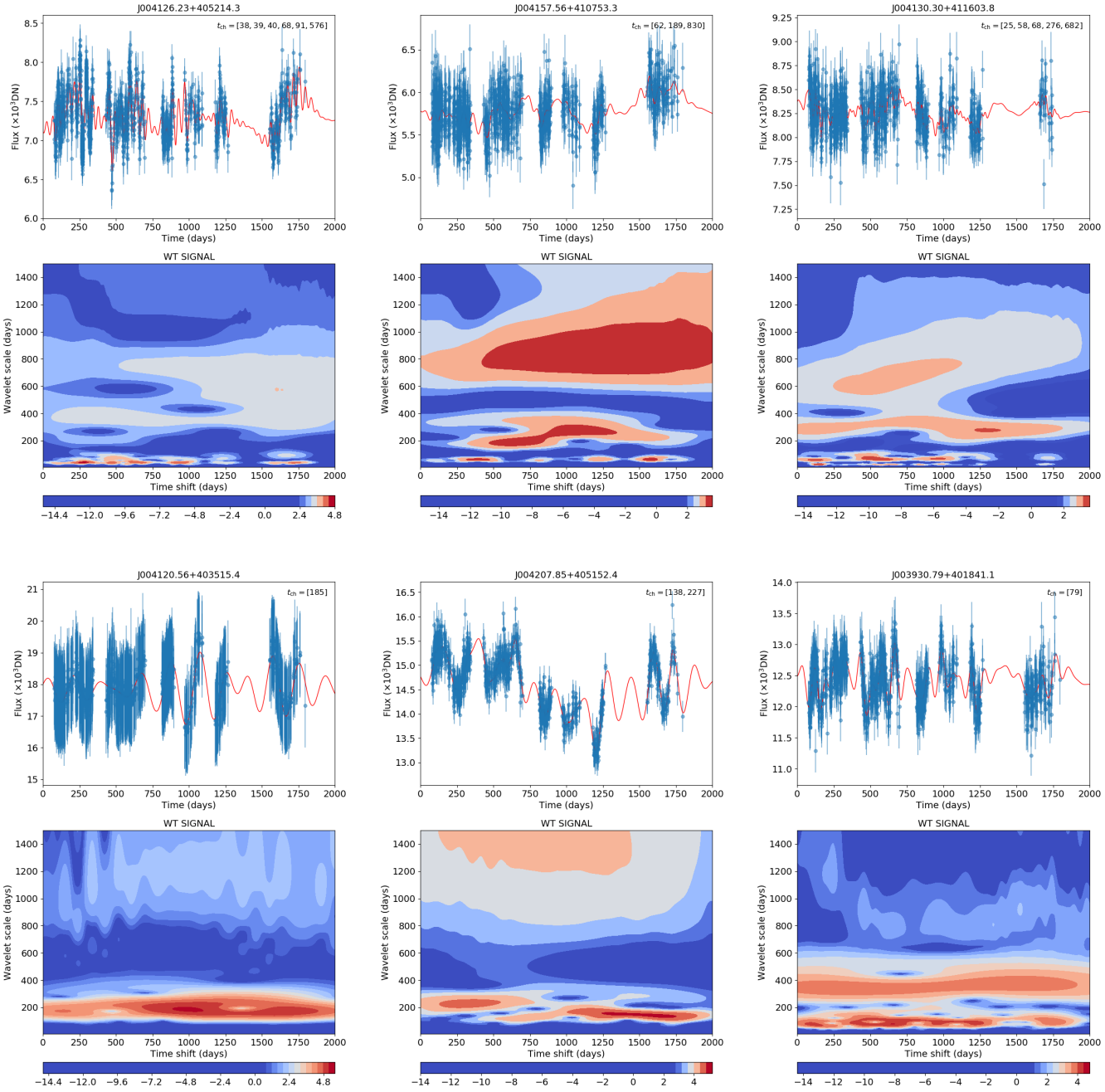

Figure 13. Contd. for AI stars (*top*) and FI stars (*bottom*).

We also perform the wavelet transform of one of the theoretical LBV light curves from the 3D simulation of Jiang et al. (2018) that shows stochastic varaibility. We find two timescales from our automated method of determining $t_{\rm ch}$, which are 4 days and 14 days with the variability amplitudes (relative to the average flux value of the light curve) being $\approx 5\%$ for both $t_{\rm ch}$ values.

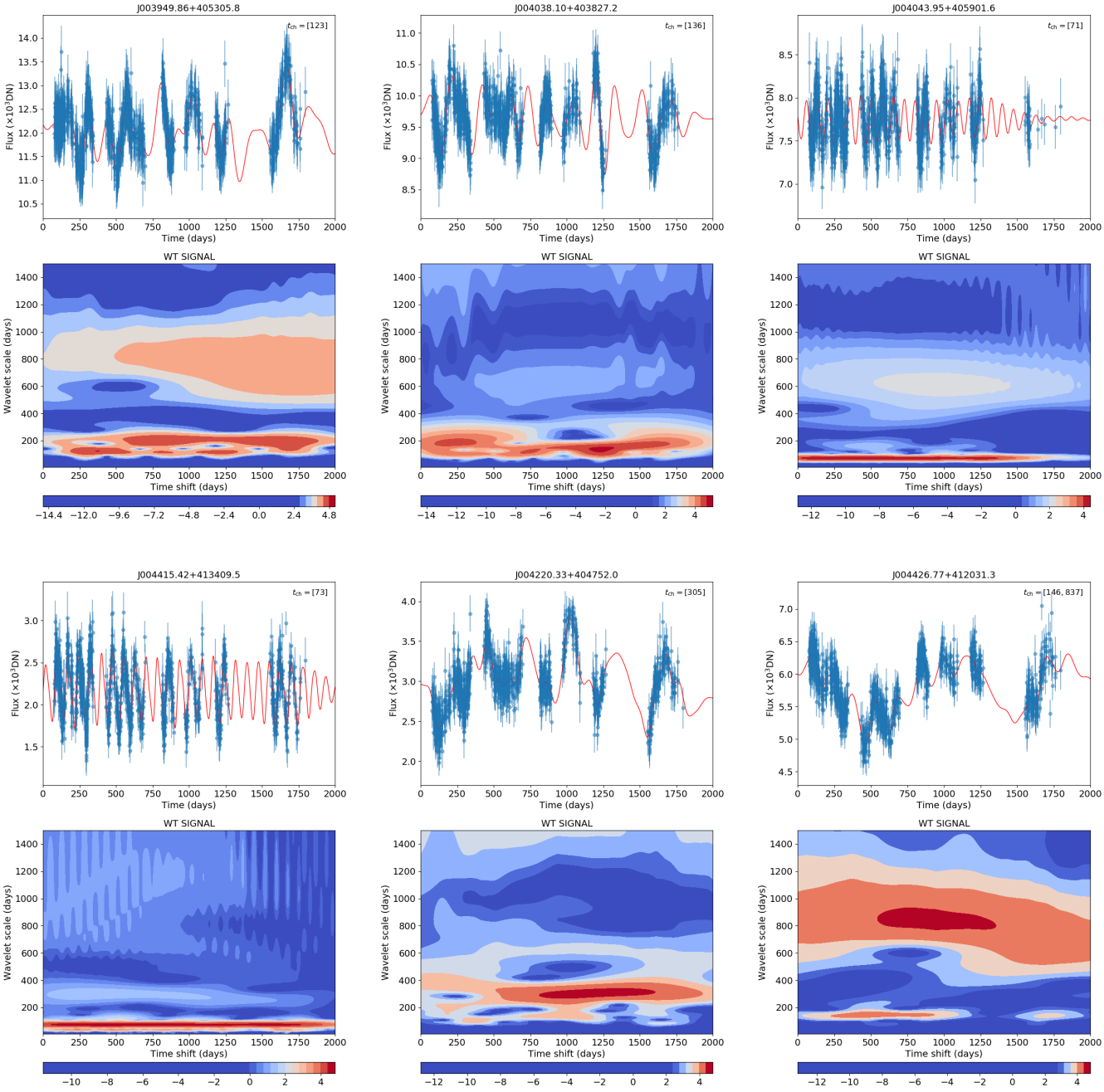

Figure 13. Contd. for GI stars (*top*) and KI stars (*bottom*).

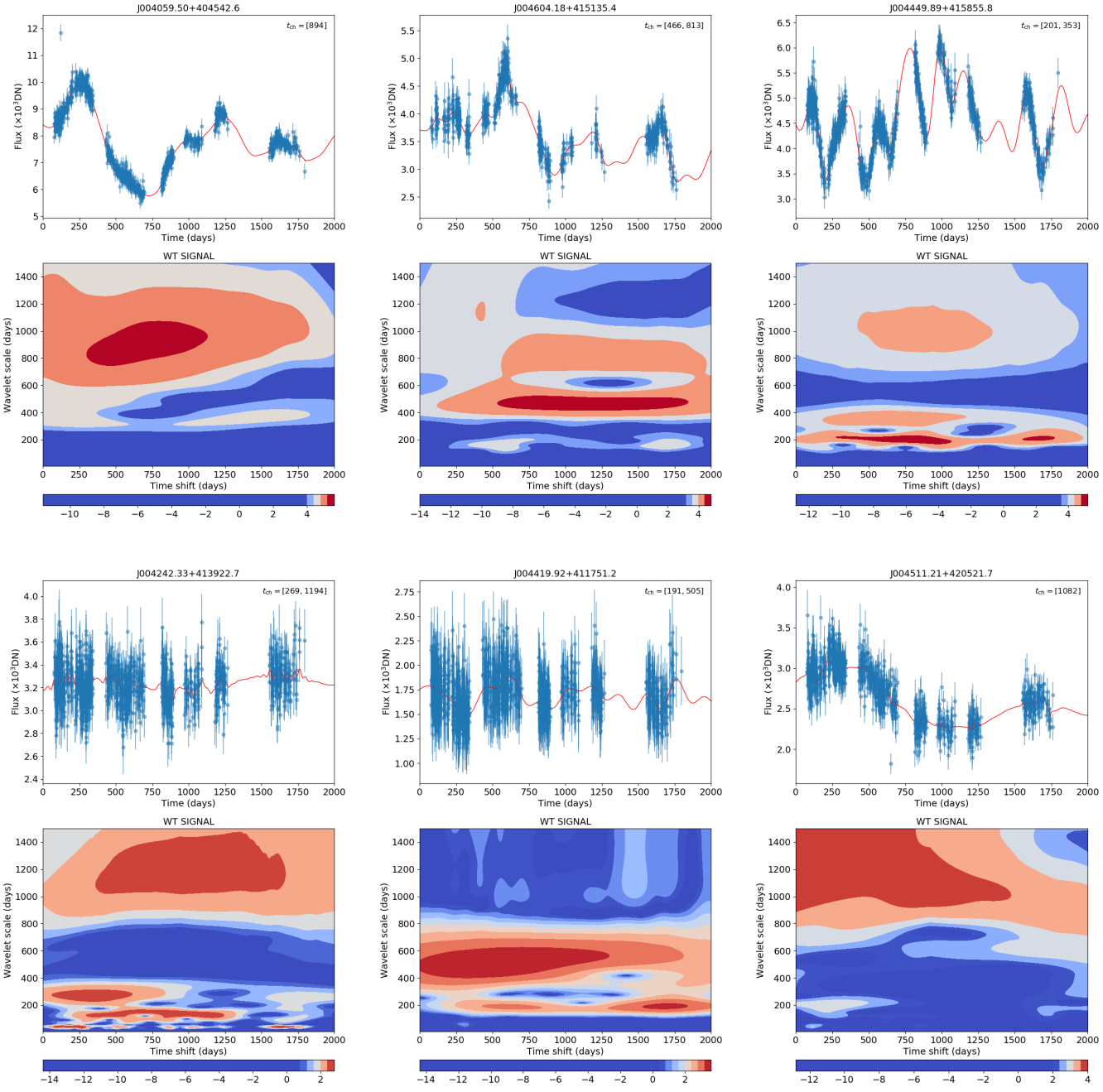

Figure 13. Contd. for M stars (*top*) and WR stars (*bottom*).

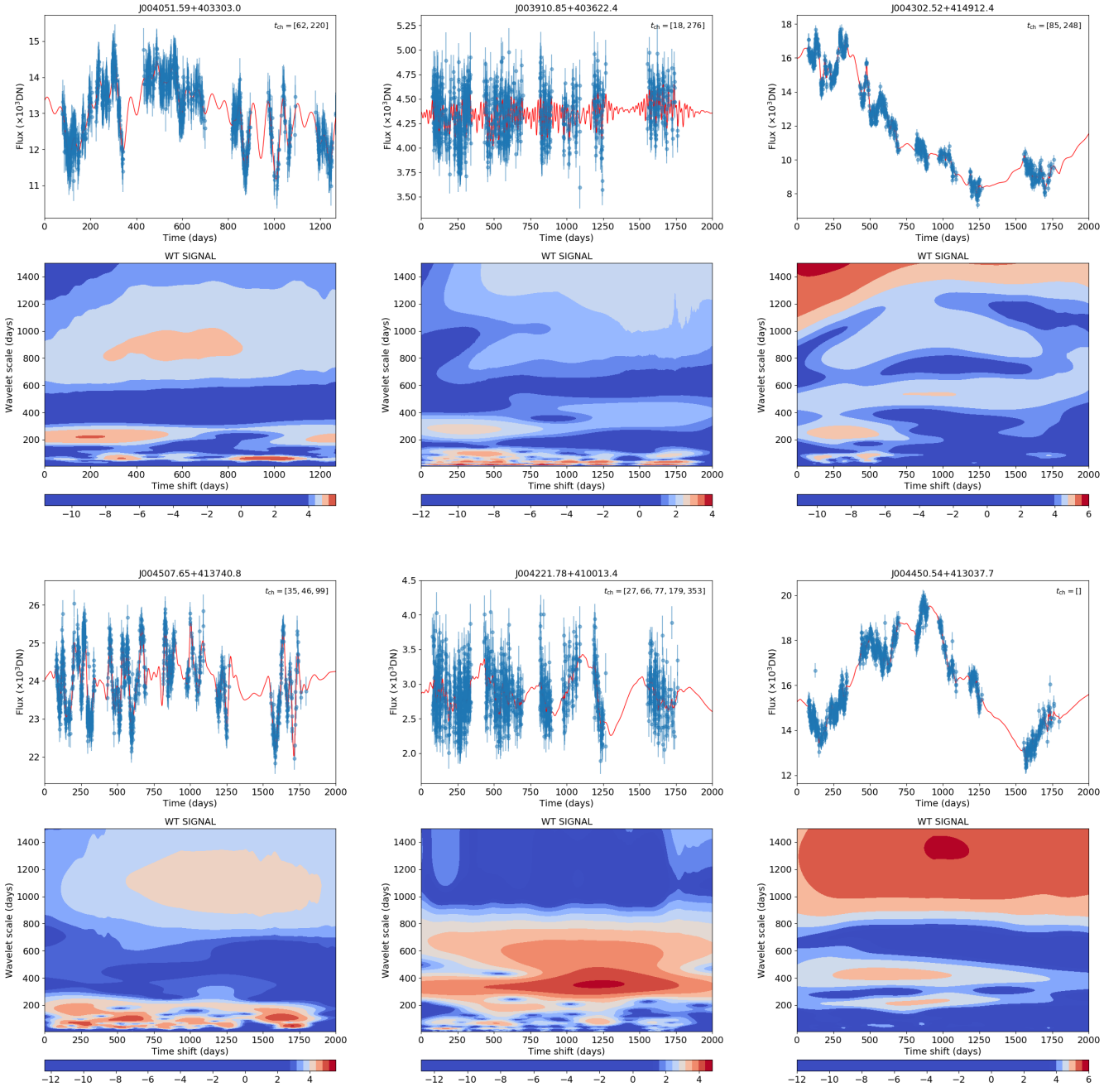

Figure 14. iPTF light curves along with the corresponding wavelet transform maps for known and candidate LBVs in M31 from MNS16. The red curve in each panel with the observed light curve shows the reconstruction (see text). The ID of the star from MNS16 is shown on top of each plot. The time axis in the light curve plots is with respect to a reference value of MJD 56000.

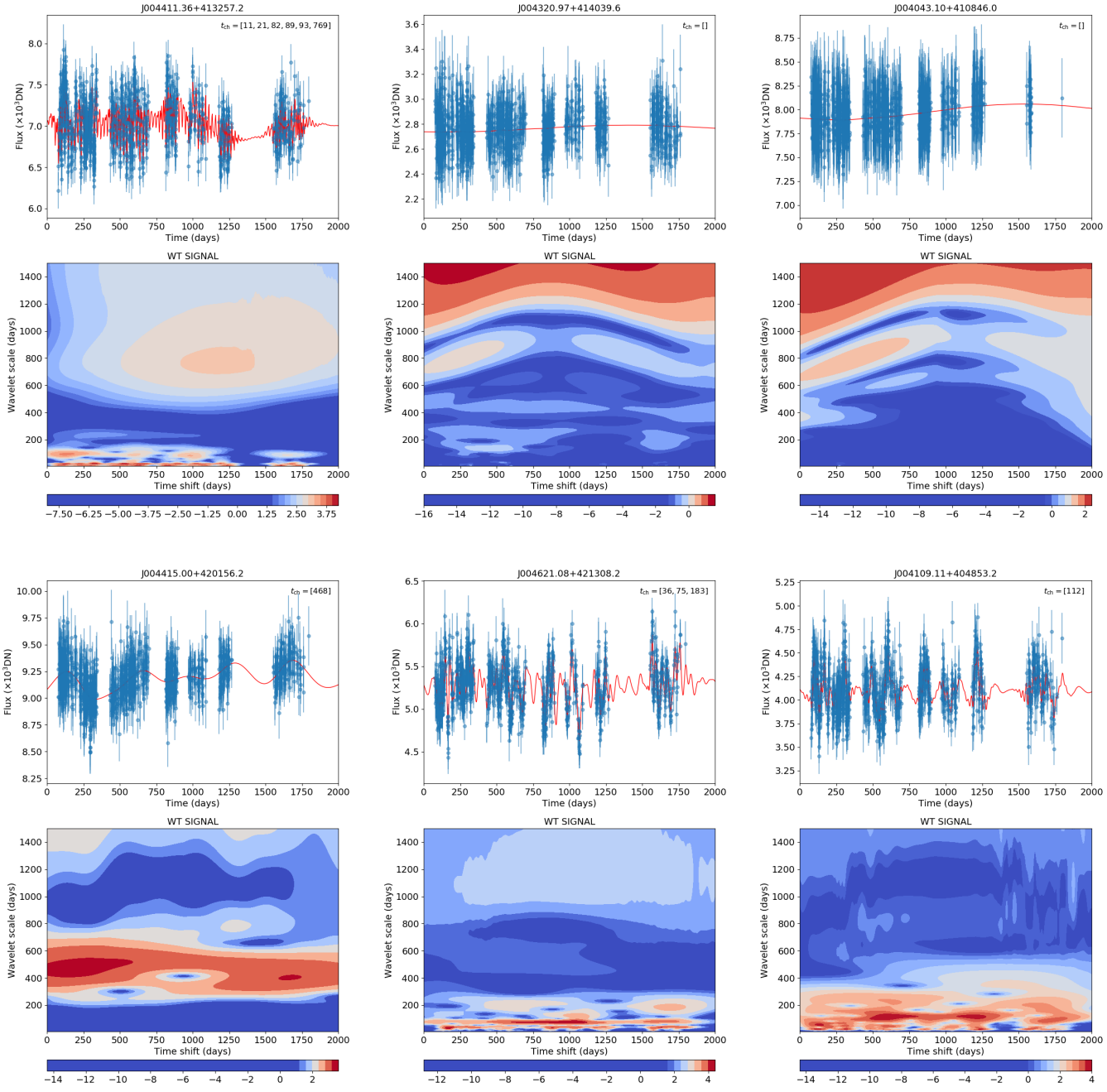

Figure 14. Contd.

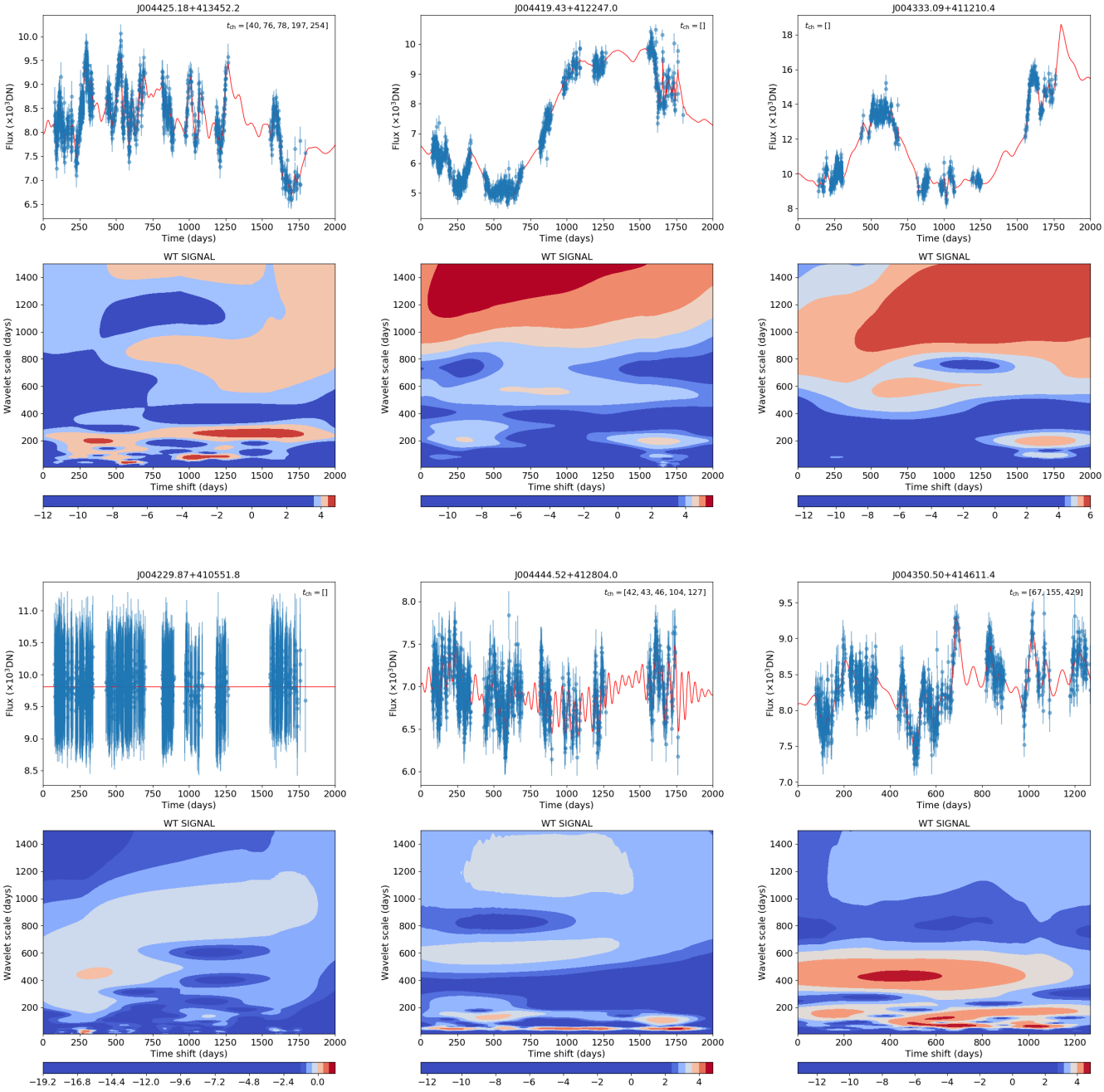

Figure 14. Contd.

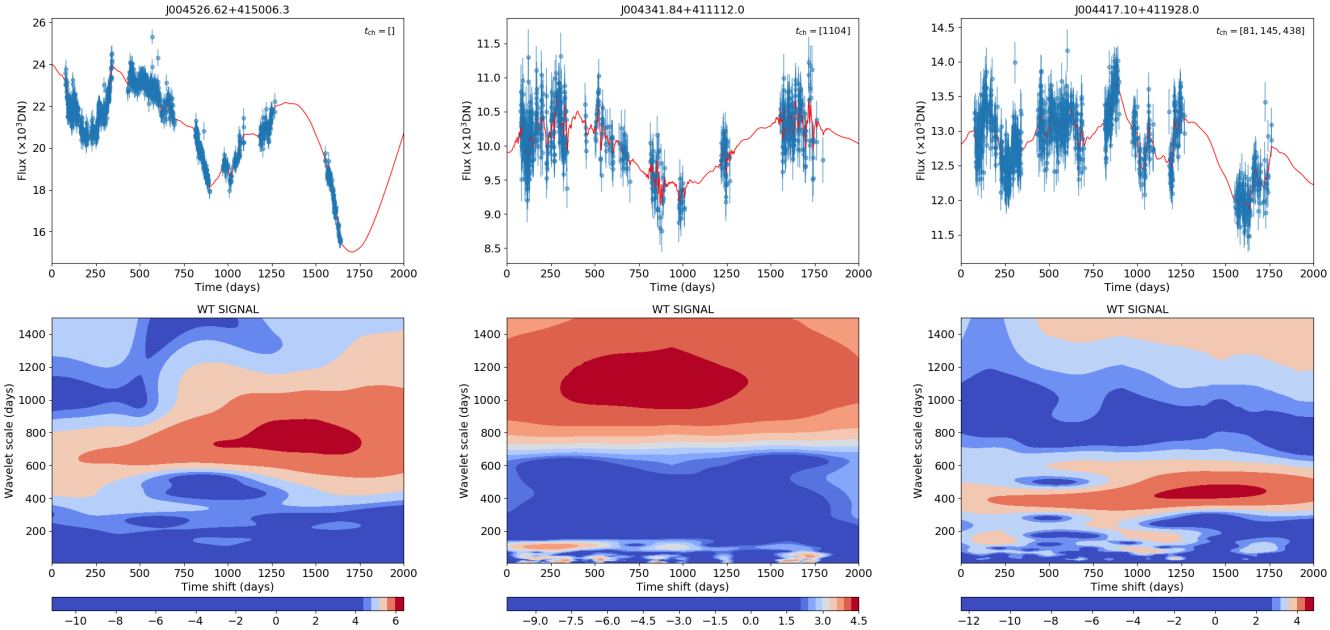

Figure 14. Contd.

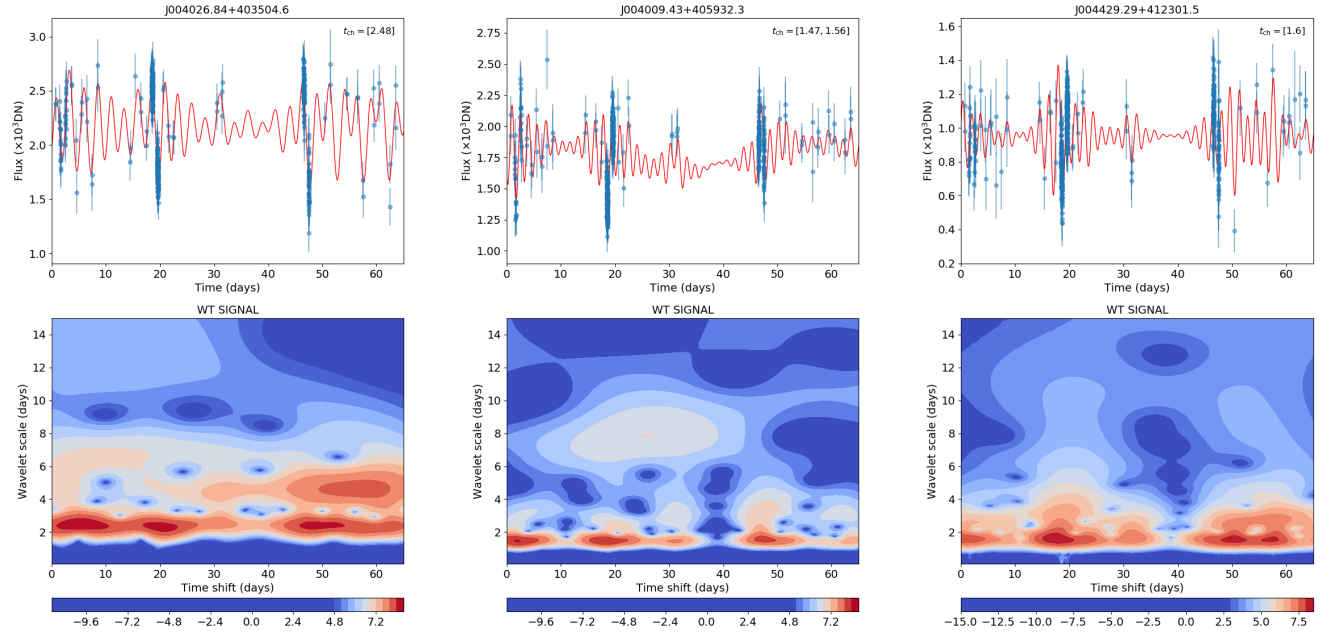

Figure 15. iPTF light curves along with the corresponding wavelet transform maps for stars with significant timescales in the high-cadence block. The ID of the star from MNS16 is shown on top of each plot and the spectral types of the stars (going from left to right) are O7+O9f:, ON9.7Iab, and O8V. The time axis in the light curve plots is with respect to a reference value of MJD 56250.621783. The wavelet transform power is shown in log scale.

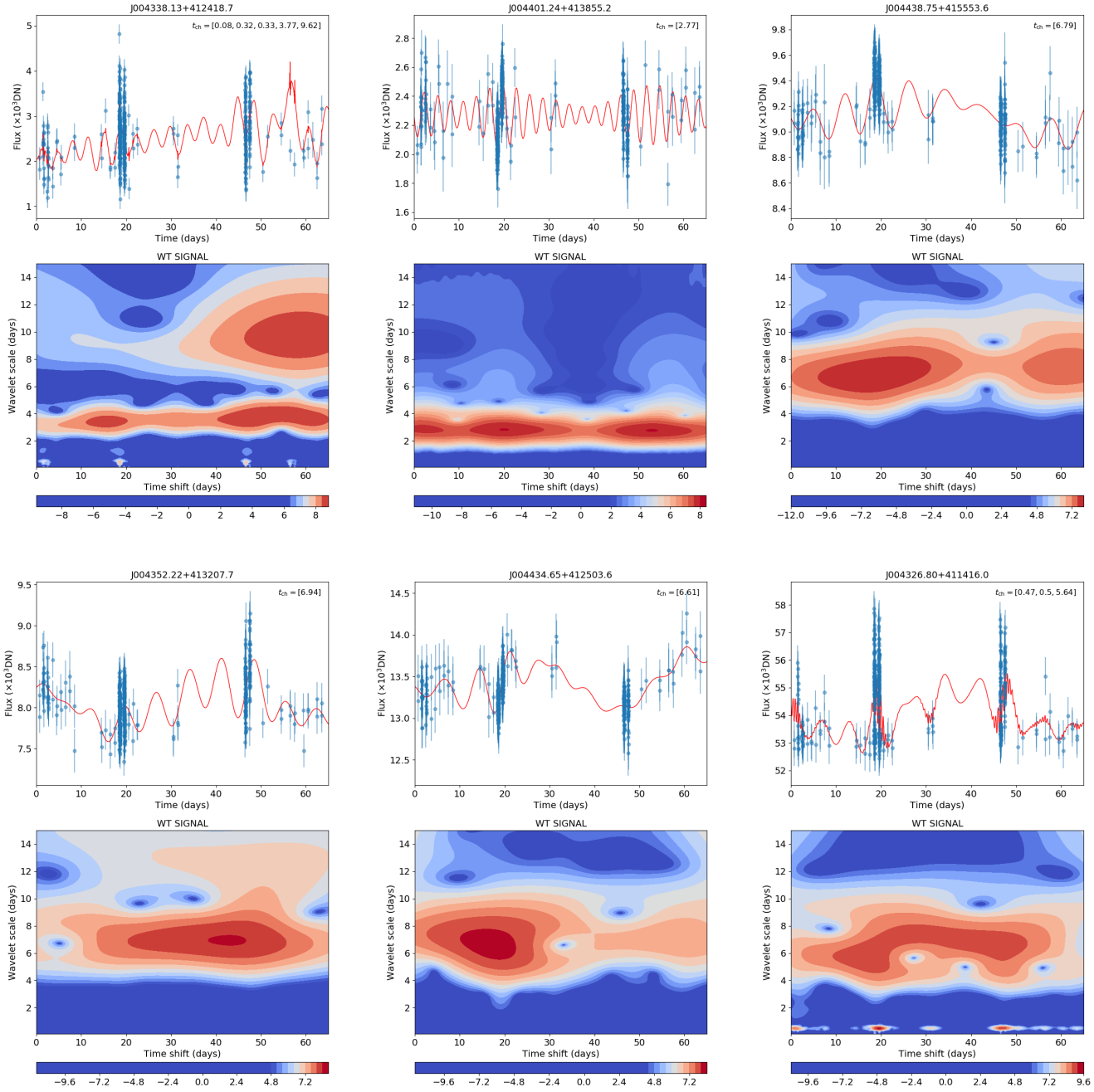

Figure 15. Contd. The spectral types of the stars (going from left to right) are B0Ia, B1.5I:, B2.5Ia, (*top* panels), and B8I, B5Ia+Neb, A:I (*bottom* panels).

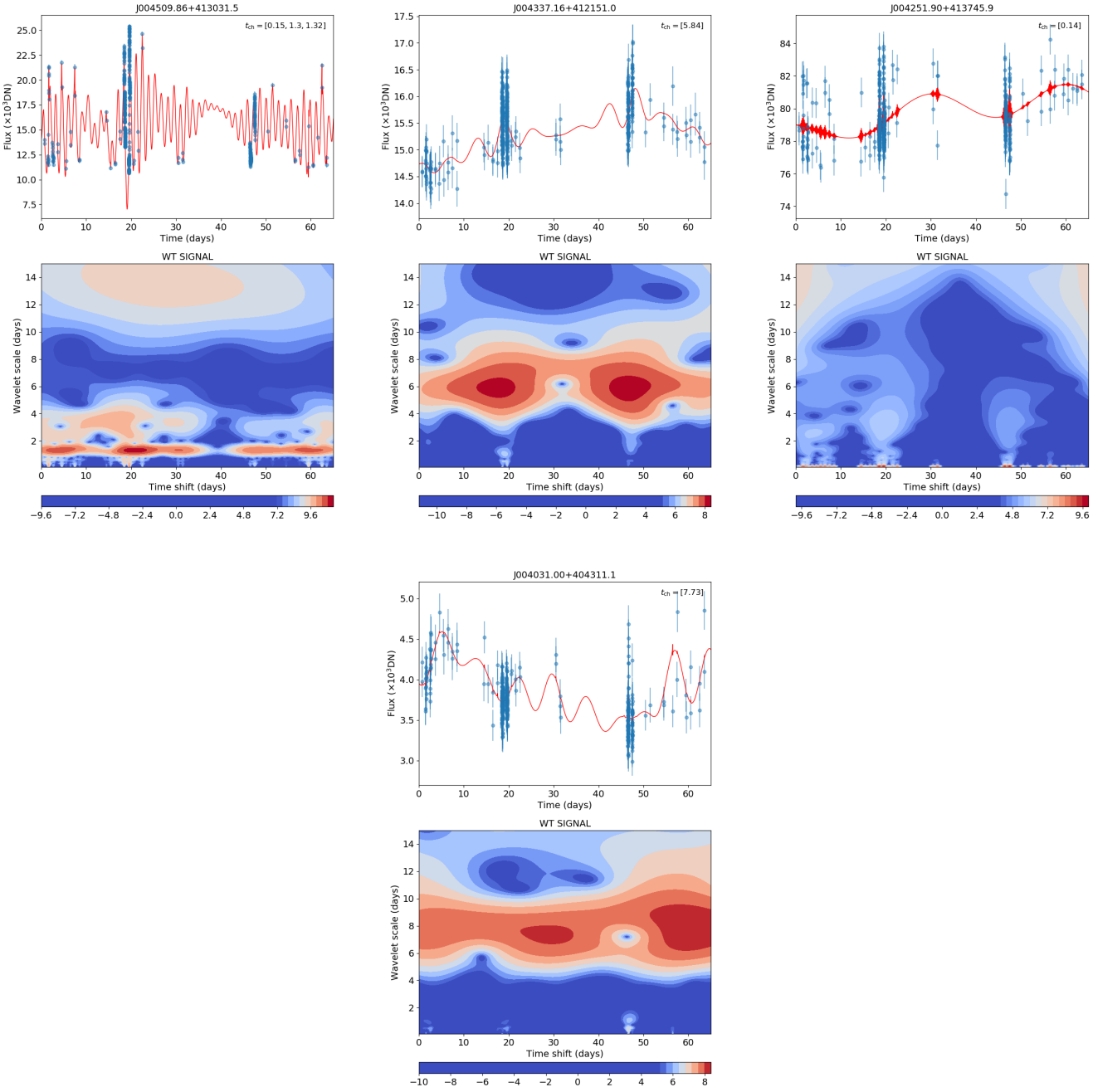

Figure 15. Contd. The spectral types of the stars are YSG: (all three *top* panels) and M0I (*bottom*).

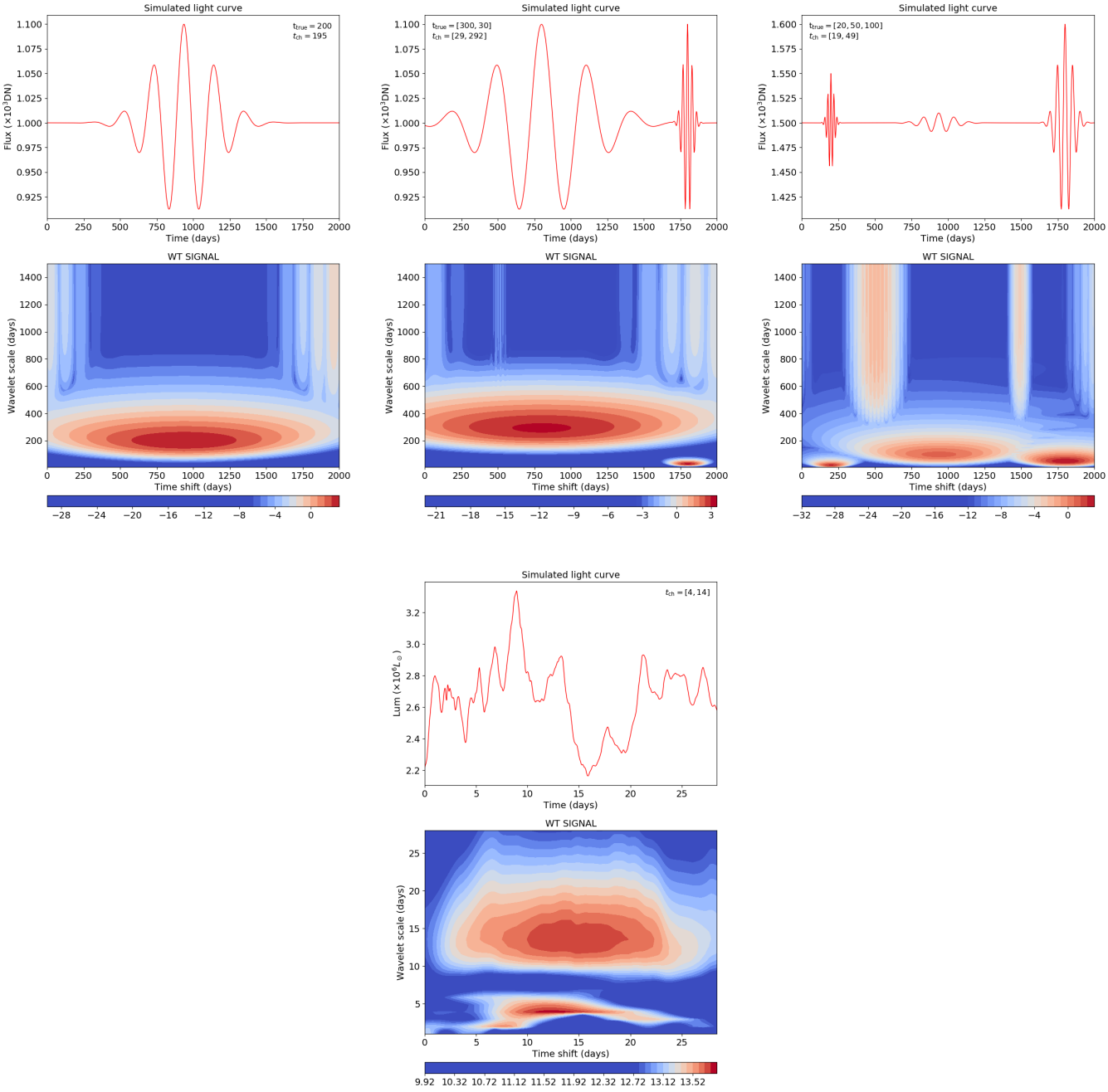

Figure 16. Simulated light curves and their corresponding wavelet transform signal (*top* three panels). The input and recovered characteristic timescales, $t_{\rm true}$ and $t_{\rm ch}$, respectively, are also indicated in the legend. The bottom panel shows the theoretical lightcurve of the LBV model with $T_{\rm eff} = 19000$ K and luminosity $\log(L/L_\odot) = 6.4$ from the 3D simulation of Jiang et al. (2018). The wavelet transform power is shown in logscale.