Astronomy & Astrophysics

Fully automated integral field spectrograph pipeline for the SEDMachine: pysedm[*]

M. Rigault[1], J. D. Neill[2], N. Blagorodnova[3], A. Dugas[2], M. Feeney[2], R. Walters[2,4], V. Brinnel[8], Y. Copin[5,6,7], C. Fremling[2], J. Nordin[8], and J. Sollerman[9]

[1] Université Clermont Auvergne, CNRS/IN2P3, Laboratoire de Physique de Clermont, 63000 Clermont-Ferrand, France
e-mail: m.rigault@ipnl.in2p3.fr
[2] Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[3] Department of Astrophysics/IMAPP, Radboud University, Nijmegen, The Netherlands
[4] Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA
[5] Université de Lyon, 69622 Lyon, France
[6] Université de Lyon 1, Villeurbanne, France
[7] CNRS/IN2P3, Institut de Physique Nucléaire de Lyon, France
[8] Institute of Physics, Humboldt-Universität zu Berlin, Newtonstr. 15, 124 89 Berlin, Germany
[9] The Oskar Klein Centre & Department of Astronomy, Stockholm University, AlbaNova 106 91 Stockholm, Sweden

ABSTRACT

Current time domain facilities are discovering hundreds of new galactic and extra-galactic transients every week. Classifying the ever-increasing number of transients is challenging, yet crucial to furthering our understanding of their nature, discovering new classes, and ensuring sample purity, for instance, for Supernova Ia cosmology. The Zwicky Transient Facility is one example of such a survey. In addition, it has a dedicated very-low resolution spectrograph, the SEDMachine, operating on the Palomar 60-inch telescope. This spectrograph's primary aim is object classification. In practice most, if not all, transients of interest brighter than ~19 mag are typed. This corresponds to approximately 10–15 targets a night. In this paper, we present a fully automated pipeline for the SEDMachine. This pipeline has been designed to be fast, robust, stable and extremely flexible. PYSEDM enables the fully automated spectral extraction of a targeted point source object in less than five minutes after the end of the exposure. The spectral color calibration is accurate at the few percent level. In the 19 weeks since PYSEDM entered production in early August of 2018, we have classified, among other objects, about 400 Type Ia supernovae and 140 Type II supernovae. We conclude that low resolution, fully automated spectrographs such as the "SEDMachine with pysedm" installed on 2-m class telescopes within the southern hemisphere could allow us to automatically and simultaneously type and obtain a redshift for most (if not all) bright transients detected by LSST within $z < 0.2$, notably potentially all Type Ia Supernovae. In comparison with the current SEDM design, this would require higher spectral resolution ($R \gtrsim 1000$) and slightly improved throughput. With this perspective in mind, pysedm is designed to easily be adaptable to any IFU-like spectrograph.

Key words. instrumentation: spectrographs – methods: observational – techniques: spectroscopic – surveys – methods: data analysis
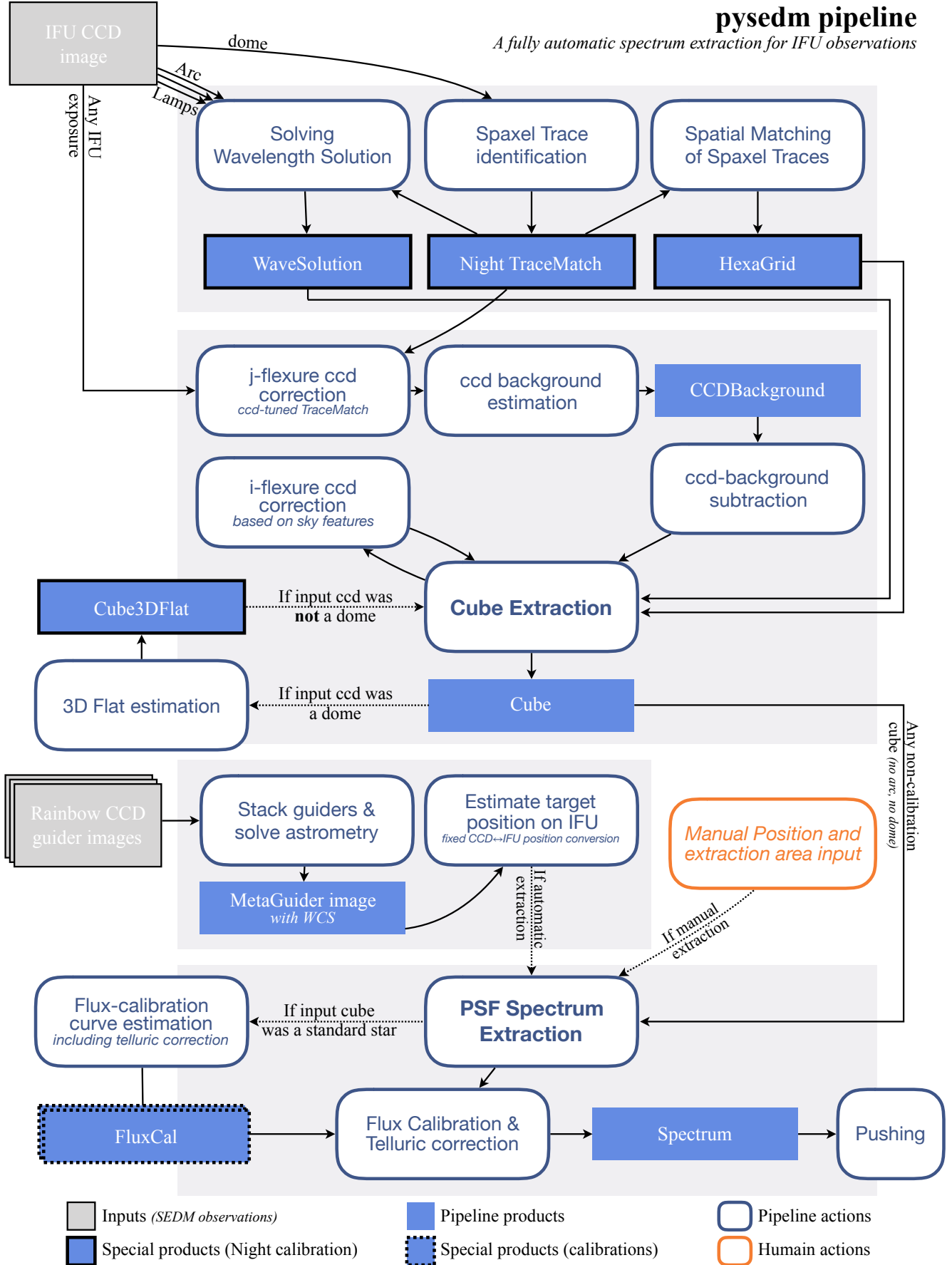
1. Introduction

Time domain astronomy, the study of transients, variables and/or moving objects, is one of the frontier fields of this decade. Modern surveys are now able to scan the entire visible sky with a daily or near-daily cadence. The Zwicky Transient Facility (ZTF, Bellm et al. 2019; Graham et al. 2019), with its 47 deg² field of view and rapid camera, is one example of such a survey. Other examples include the Catalina Real-Time Transient Survey (Drake et al. 2009), PanSTARRS-1 (Kaiser et al. 2002), ASAS-SN (Shappee et al. 2014) and ATLAS (Tonry et al. 2018). While scanning the sky every night with a typical 5σ magnitude limit of 20.5, ZTF detects of order 10^5 variations in the sky between a new observation and a past reference frame, generating an alert for each. While most of these alerts are glitches or already known variable sources, about $O(10^2)$ of them are transients of interest. These may be, for example, supernovae or any kind of galactic or extra-galactic explosion. About 10% of them are new transients that have just exploded, or just become bright enough to be detected. Soon the Large Synoptic Survey Telescope (LSST, LSST Science Collaboration 2009), with a magnitude limit of 27.5 in r-band, will detect ten times more transients than ZTF, resulting in hundreds of new extra-galactic events every night.

In that context, classifying the ever-increasing number of new supernovae (or equivalent) is a technical challenge. Yet a good classification is crucial to furthering our understanding of the nature of these transients, to investigate their rates and to detect new classes of objects. For cosmology, an accurate classification is required to ensure that the Hubble diagram is not contaminated by non-Type Ia supernovae. A major effort is underway to develop photometric typing (e.g., Sako et al. 2008; Ishida & de Souza 2013; Lochner et al. 2016; Möller et al. 2016, and references therein). However, these methods have not yet proven to be sufficient, notably in cosmology, where non-Ia contamination could significantly bias derivations of the cosmological parameters (e.g., Jones et al. 2017). In addition, without a spectral identification, new object families could be left undiscovered.

We first presented the SEDMachine (SEDM) in Blagorodnova et al. (2018). SEDM is a very low-resolution ($R \sim 100$) fully-filled integral field spectrograph with a ~28 × 28 arcsec² field-of-view,
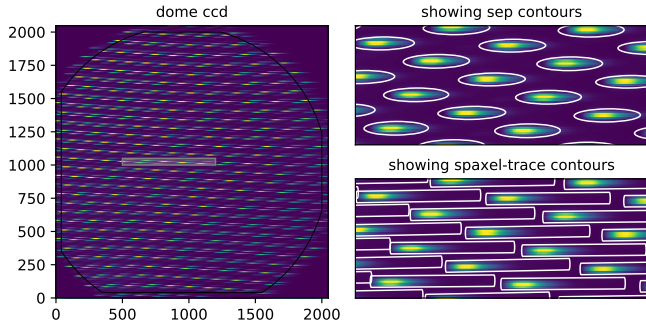
Fig. 1. pysedm pipeline.

Fig. 2. Extracting the trace contours from a SEDM dome exposure. *Left*: CCD image of a dome exposure. The gray rectangle shows the zoom area displayed on the *right*. *Upper right*: sep ellipses overplotted on the CCD dome image. *Lower right*: spaxel-trace contours overplotted on the CCD dome image. For the SEDM, spaxel-trace contours are defined as tilted rectangles. The rotation angle of the centroid position are derived from the sep ellipses.

subdivided into ~1300 hexagonal spaxels, with a projected diameter of 0.75 arcsec each, by a specially designed micro-lens array (MLA). SEDM also has a 13×13 arcsec^2 field-of-view photometric channel that simultaneously observes the sky in four different filters by subdividing the field-of-view into quadrants that are filtered, one each in the u, g, r, and i bands. This channel is mainly used for positioning and guiding, and, occasionally for follow-up purposes. SEDM has been designed and built to type transients discovered by ZTF and with pysedm we typically acquire 15 transient spectra a night. The IFU allows us to perform both host and sky background subtraction using a single exposure of a given transient. Since the beginning of this survey in early June 2018, ZTF has discovered and classified over 1000 new extra-galactic transients, most of them Type Ia supernovae with $z < 0.15$ Feindt et al. (2019). The SEDM demonstrates that typing a large fraction (if not all brighter than a certain magnitude) of detected nearby transients is doable with dedicated facilities, typically $\lesssim$20 mag if installed on 1–2 m class telescopes. For such a facility to be efficient, however, two key elements are necessary: full automation and high reliability. While the selection of transients sent to the SEDM is currently only partially automated, since early August 2018 the new SEDM pipeline, pysedm, has been in production and has extracted all submitted transient spectra automatically.

In this paper we present pysedm. This software is meant to be a generic pipeline that could easily be adapted to any further SEDM-like instrument. The code is public and written in Python[1]. It is not designed to optimize the spectrophotometry, but rather to be fast, robust, fully automated, and easily customizable.

The pipeline is illustrated in Fig. 1. In Sect. 2 we explain the cube extraction from the raw observations of the nightly calibration solutions including the trace identification (Sect. 2.1.1), and the wavelength solution (Sect. 2.1.2). Then, we present in Sect. 3 the spectral extraction from the 3D cubes. An overall pipeline performance and some results are presented in Sect. 4. We conclude in Sect. 5.

2. Extracting a 3D cube

In integral field spectroscopy, extracting a 3D-cube from a 2D-image requires the identification of three calibration elements: (1) where spaxel light is dispersed onto the CCD – forming a

"trace", (2) the correspondence between the location of a trace on the CCD and the position of the spaxel in the focal plane and (3) the spectral dispersion mapping that converts each CCD-pixel into actual wavelengths. Throughout the paper, we refer to the 3D "x, y, λ" object as a "cube". This contains a spectrum (λ) per sky position (x, y), integrated over a spatial element called a spaxel. In this generic appellation, a spaxel might have any shape and be located anywhere (Kissler-Patig et al. 2004). Therefore, a 3D "cube" does not necessarily has a square foot-print and is not necessarily fully-filled. In our case, SEDM has hexagonal spaxels forming a fully-filled tessellated roughly circular field-of-view (see Fig. 4).

In the implementation of pysedm for the SEDM, the three aforementioned calibration elements are derived once for the entire night based on dome and arc-lamp observations taken during late afternoon. During nighttime observations, instrumental flexure is solved for using night sky spaxels, and each image is aligned to the nightly calibration solution. This approach enables us to maximize the time dedicated to science exposures, as no further calibration observations are required. In addition, since the calibration steps are made during daytime, the pysedm pipeline can extract an automatic target spectrum at night a few minutes min after the end of the exposure.

In this section we detail the cube extraction mechanism. We start by presenting the night calibration solutions in Sect. 2.1, then in Sect. 2.2 we explain how to use them to extract the 3D cubes.

2.1. Night calibration

Here we present the details of the algorithmic solutions developed in pysedm to build the calibrations representing the three aforementioned elements: trace matching, spatial identification, and the wavelength solution. An illustration of a raw SEDM CCD image is shown in Fig. 2; here a flat dome. We see that spaxel traces are dispersed nearly horizontally, with bluer wavelengths toward the right of the traces.

2.1.1. Trace matching

– Input: `dome.fits`;
– Output: `TraceMatch.fits`.

To identify each individual spaxel trace, we ran the EXTRACT method of sep[2] (Barbary 2016) – the Python implementation of SEXTRACTOR (Bertin & Arnouts 1996) – on a dome CCD-exposure. The sep function defines an ellipse contour on dispersed dome light for each spaxel; see Fig. 2. In the pysedm implementation for the SEDM, spaxel traces are defined as the rectangles formed by taking a 2.4 pixels region above and below the spaxel ellipse centroid – so 5.8 pixels in total – and starting 70 x-pixels to the left of the ellipse centroid and ending 230 x-pixels to the right; the dome light times the instrument response been mostly red.

The 2.4 pixel width has been selected to encapsulate 95.5% ($\pm 2\sigma$) of the spaxel light given the SEDM cross dispersion. As shown in Fig. 2, the compactness of spaxel traces in our instrument does not allow for significantly larger width. The trace extent of 70 pixels to the left and 230 pixels to the right has been chosen to be the smallest length that always encapsulates wavelengths from 3500 to 9500 Å, after wavelength solution conversion (see Sect. 2.1.2).

---

[1] https://github.com/MickaelRigault/pysedm

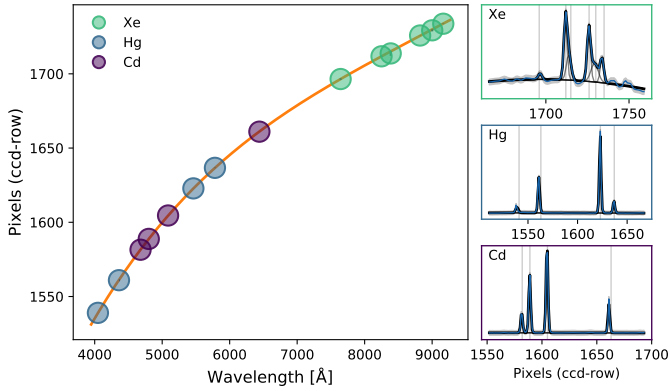[2] http://github.com/kbarbary/sep; version used: 1.0.1.

Fig. 3. Wavelength solution fit for spaxel #501. *Left*: fitted centroids of Hg, Cd and Xe (see legend) arc lamp emission lines (in pixel-row from the right) as a function of their expected wavelengths. The orange line is the best fitted fifth order polynomial enabling the conversion of wavelength to pixel and conversely. *Right*: blue curves represent the arc lamp spectra (in flux per pixel-row from the right). The black lines are the best fitted "Gaussians + third-order polynomial-continuum". Vertical thin gray lines are the initial guess positions. See details in Sect. 2.1.2.

Once a spaxel trace is defined, we can create its corresponding 2D CCD weight-mask defined such that each pixel corresponds to the fraction of its area that lies within the trace contours; zero for outside, one for fully inside, and intermediate values for pixels on an edge. This fraction is measured using SHAPELY[3].

Given this 2D mask and because the spectral dispersion of the SEDM is almost horizontal, one can extract a spaxel spectrum in counts per pixel (named "pixel-spectrum") by summing the product "CCD-image × spaxel weight-mask" along the vertical-axis ("*j*-axis"). For convenience, when extracting a pixel-spectrum we count the pixels from right to left such that increasing pixels correspond to increasing wavelengths.

2.1.2. Wavelength solution

– Input:
  - Hg.fits, Xe.fits, Cd.fits;
  - TraceMatch.fits (see Sect. 2.1.1);
– Output: WaveSolution.fits.

A wavelength solution is derived for each spaxel independently based on arc-lamp observations. We used three lamps for the SEDM wavelength calibration (wavelength in vacuum); see Fig. 3:

Hg: A mercury lamp (Hg) that has four strong emission lines at 4047.7, 4359.6, 5462.3 and 5781.7 Å (a blend of 5771.2 and 5792.3).

Cd: A cadmium lamp (Cd) that has four strong emission lines at 4679.3, 4801.3, 5087.2 and 6440.2 Å.

Xe: A xenon lamp (Xe) that has six emission lines at 7644.1, 8250.1 (a blend of 8233.90 and 8282.39), 8386.2 (a blend of 8349.1 and 8411.0), 8821.8, 9001.3 (a blend of 8954.7 and 9047.9) and 9165.1 Å. The bluer Xe emission line is faint, but we decided to use it nonetheless to minimize the gap around 7500 Å; see Fig. 3.

For a given spaxel, we extracted the corresponding pixel-spectrum (see Sect. 2.1.1) for each of the three arc lamps and independently fit them with a combination of Gaussian lines plus
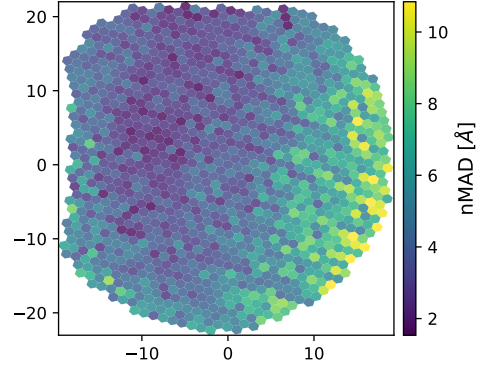
---

[3] https://pypi.org/project/Shapely; version used: 1.6.3.

Fig. 4. Illustration of the precision of the wavesolution procedure. The color of each spaxel corresponds to the normalized median average distance (nMAD) of measured arc lines around the wavelength solution. As discussed in the text, the meaningful information is the relative precision between spaxels: our wavelength solution is better toward the center of the IFU, where the signal-to-noise ratio is higher, and science targets are designed to be located.

a third-order polynomial continuum. The Xe fit uses six Gaussian lines while the Cd and Hg fits only have four. In a second iteration, we jointly fit the bijective relation between the 14 centroids (in pixels) as a function of their expected wavelengths with a fifth-order polynomial. The best fitted polynomial is recorded and defines the wavelength solution for the given spaxels. An illustration of the procedure is shown in Fig. 3.

By construction, each spaxel wavelength solution is independent. As a validation plot for this calibration step, we record the normalized median average distance (nMAD) around the wavelength solution for each spaxel; a typical example is shown in Fig. 4. This scatter plot contains both Gaussian line fit noise and potential inaccuracy in the "true wavelength" definition for the arc lamps, especially for the blended lines. Hence, while the absolute nMAD value might not be so relevant, the relative precision is meaningful and we see in Fig. 4 that the wavelength solution is more precise in the center of the MLA than on its right edge (see details on spatial solution in Sect. 2.1.3). If a wavelength solution fails for a given spaxel, its nMAD would be much larger than a few tens of ängstroms. While, so far, this has not happened as long as the pipeline has been in production, such a figure would enable us to identify any potential problems with the wavelength solution calibration step. A test of the wavelength calibration precision is presented in Sect. 3.4.

2.1.3. Spatial identification

– Input: TraceMatch.fits;
– Output: HexaGrid.fits.

The SEDM has a hexagonal MLA located in the focal plane of the Palomar 60″ (P60) telescope. The SEDM therefore creates a fully-filled hexagonal x, y, λ cube. The task of the spatial identification step is to recover this hexagonal structure given the spaxel dispersion trace location on the CCD. This is done in eight steps.

First, we used the spaxel-trace centroid to define spaxel reference positions on the CCD. We use the trace-id from TraceMatch.fits to identify the centroids. Second, we integrated the centroids into a k-d tree and associate each spaxel with its six nearest neighbors. Given the hexagonal structure of the MLA, these are the six contiguous lenses.

ZTF18abqlpgq (aka SN2018fsf)

SDSS finder chart

SEDmachine reconstructed image | $\lambda \in [5000, 8000]\,\text{Å}$
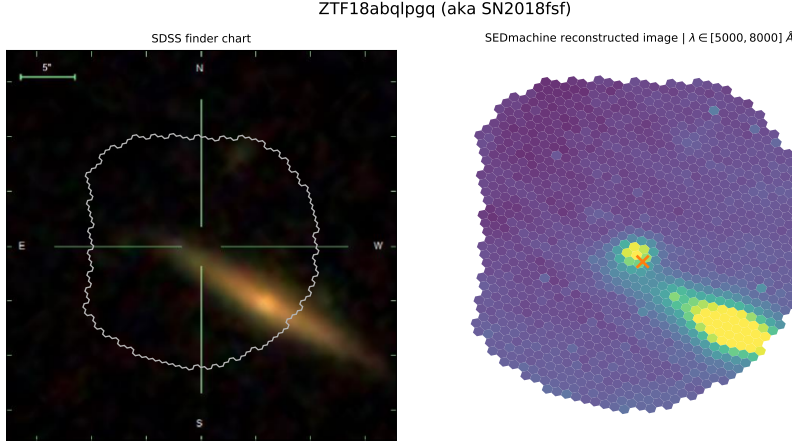

Fig. 5. Illustration of the SEDM field of view; ZTF18abqlpgq. *Left*: RGB SDSS image centered at the target (ZTF18abqlpgq) location. The gray line represents the SEDM footprint for this target. *Right*: integrated 2D image of the corresponding SEDM 3D-cube. The orange cross shows the expected target location used as first guess for the automatic PSF extraction. We clearly see the point source near its expected position as well as the host galaxy.

Third, we started the hexagonal reconstruction around a "reference spaxel", set at the beginning to the most central centroid on the CCD. We then picked the nearest of the six neighbors to form a "spaxel-pair". By construction, this pair only has two shared neighbors, except if this pair is located at the edge of the MLA in which case it only has one.

Fourth, we selected the closest of the two shared neighbors of the spaxel-pair to form a triangle. This first such triangle defines the $\{q, r\}$ coordinates of the hexagonal structure. Fifth, we formed a new "spaxel-pair" by pairing the reference spaxel with the third spaxel of the triangle. By construction, this pair only has at maximum one unknown neighbor, since we already know they have the "second-spaxel" in common. If it exists, meaning if the pair is not at the edge, we recorded the $\{q, r\}$ coordinates of this fourth centroid. At this stage, the reference spaxel and the third spaxel connect with three of their neighbors, while the fourth and the second connect to only two.

Sixth, we repeated the fifth step updating the pairing with the reference spaxel to the latest associated spaxel. After three iterations we can connect all the six reference spaxel neighbors, and each of these are already connected to three of their six neighbors.

Seventh, we repeated the fifth and sixth steps by randomly selecting one of the six neighbors of the reference spaxel to become the new reference. We continued until all spaxel centroids have been assigned.

Finally, a rotation of 263° is applied to the reconstructed hexagonal structure such that north is up and east is to the left. A reconstructed image is shown in Fig. 5, along with its corresponding optical image. We recover the galaxy visible in the SDSS finding chart in the MLA. The MLA also contains the transient, here ZTF18abqlpgq, aka SN 2018fsf, a $z = .095$ Type Ia supernova near maximum light.

2.2. CCD to cube

– Input:
  • `crr_target_file.fits`,
  • `TraceMatch.fits`,
  • `HexaGrid.fits`,
  • `WaveSolution.fits`.
– Output: `e3d_target_file.fits`.

The cube extraction makes use of the night calibration produced during the daytime to extract the 3D cube from a given CCD image. In addition, to account for instrumental flexure, the cube extraction has two independent flexure-correction steps, one for the trace location and one for the wavelength solution. The spatial solution is fixed. The cube extraction is therefore done as

follows, with each step detailed further in the following subsections for interested readers:

1. Trace-position flexure correction: vertically move together the trace locations to optimize the trace alignment.
2. CCD background removal: mask out regions contained within the flexure-corrected traces and construct a background CCD-image based solely on the out-of-trace pixels. This background is stored and removed from the CCD image.
3. Cube extraction: For each flexure-corrected trace, extract its corresponding pixel-spectrum and convert it into a wavelength-spectrum using the night wavelength solution. Once all the spaxels are created this way, combine them into a cube using the nightly spatial solution.
4. Wavelength flexure correction: Measure the effective wavelength position of sky and telluric lines and convert the observed wavelength shift $\Delta\lambda$ into a horizontal Δi pixel shift.
5. Flexure corrected cube extraction: Repeat step 3, applying the Δi pixel shift identified at step 4 when converting pixel into wavelength.
6. Flat field the Cube: correct the relative spaxel responses using the cube-flat made from the dome exposure.

Once done, the cube is stored as a fits file. The file structure is a simplified version of the flexible euro3d format (Kissler-Patig et al. 2004). We invite the reader to use the PYIFU python library to open these cubes. A cube fits file contains five entries:

The primary image contains an $M \times N$ array, where M is the number of spaxels and N is the length of each spectrum. The corresponding wavelength array can be reconstructed from the header, using the starting wavelength and the wavelength step recorded in the primary header as CDELT1 and CRVAL1, respectively.

The second image, "VARIANCE", contains an $M \times N$ data array corresponding to the spaxel spectral variance.

The third image, "MAPPING", has a $2 \times M$ data array. This records the $\{x, y\}$ coordinates of the M spaxel centers.

The fourth image, "SPAX_ID", has an M-length array where the spaxel "id" is stored. This "id" corresponds to the "trace-id" and enables the user to easily map a given spaxel from a 3D-cube to its corresponding dispersion trace on the CCD 2D-image.

The fifth image, "SPAX_VERT", contains the spaxel vertices.

As in euro3d, the format adopted in PYIFU enables work with cubes of any structure: spaxel positions in the focal plane are recorded – and having spaxels of any shape – their vertices are stored. However, our simplified version of euro3d enforces all spaxels to share the same shape and their spectra to share the same wavelength binning.
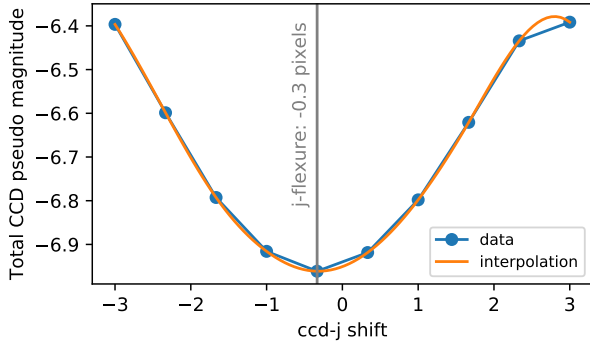
Fig. 6. Illustration of the vertical trace flexure procedure; here a science observation of ZTF18abqlpgq. Blue markers show the pseudo-magnitude computed inside the spaxel traces, the lower the brighter. The orange curve is a cubic spline interpolator. The gray vertical bar shows the minimum of the cubic interpolation. In this illustrative case, the dome-based trace mask must be shifted by −0.3 pixels to account for vertical flexure.

2.2.1. Trace position flexure correction

The spaxel trace positions are predefined for the entire night based on dome observations (see Sect. 2.1.1). However, while the telescope is moving and pointing toward a science target, it is expected that the instrument alignment will change slightly. Therefore, the original trace position might not be optimal. Given the SEDM configuration, vertical misalignment will translate into signal-to-noise ratio reduction, as a fraction of the light will not be included inside the trace – and CCD-background bias (see Sect. 2.2.2) – where we have assumed that light outside the traces are pure background. On the other hand, horizontal flexure will cause wavelength miscalibration. While this might not be optimal, we have decided in PYSEDM to treat both flexure effects independently. See Sect. 2.2.4 for correction of the horizontal displacement.

To measure the vertical trace displacement for a given exposure, we measure the total flux T_f contained within all spaxel traces while vertically moving them from −3 to +3 pixels around the default position in ten regular steps. Each time, a pseudo-magnitude is defined as $-2.5 \times \log(T_f)$. For an illustrative observation of ZTF18abqlpgq, Fig. 6 shows the evolution of this pseudo-magnitude as a function of the vertical shift applied. We interpolated the ten measured pseudo-magnitudes with a cubic spline to find the vertical shift minimizing the pseudo-magnitude. In this example the shift is −0.3 pixels; shifts rarely go beyond ±1 pixel. In the rest of the cube extraction procedure, the nightly TraceMask solution is vertically shifted by +0.3 pixels to counterbalance the vertical instrument flexure.

2.2.2. CCD background subtraction

After the trace flexure correction, light outside the traces may be considered to be diffuse. In order to remove this background light from the actual spaxel signal, we build a CCD-background that will be removed from the CCD image. This background is constructed by fitting every ten columns, a fifth-order Legendre polynomial on pure-background pixels, meaning those in none of the spaxel traces. The actual background CCD-image is built by concatenating these fitting polynomials and interpolating between them with a horizontal Gaussian kernel with a width of five pixels.

2.2.3. Cube extraction

For each spaxel, we first created a pixel-spectrum (see definition in Sect. 2.1.1) using the flexure-corrected spaxel trace mask
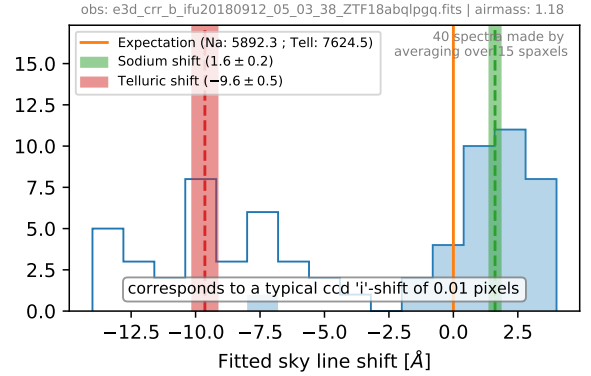
Fig. 7. Illustration of the horizontal wavelength flexure procedure; here a science observation of ZTF18abqlpgq. Histograms show the measured wavelength shift between expected and measured sky sodium lines (blue) and telluric absorption (white). The measurements are made on the 600 faintest spaxels (∼ half of the MLA) into 40 spectra averaging 15 spaxels. The orange vertical bar shows the average shift in wavelength. The text indicates the conversion in terms of horizontal pixels. In the illustrative case there is barely any wavelength flexure correction, which rarely exceeds a fraction of a pixel.

on the background-subtracted CCD-image. We converted pixels to wavelengths using the wavelength solution to create the wavelength-spectrum of this spaxel. This wavelength solution is the nightly wavelength solution onto which we can apply a pixel shift to account for wavelength flexure (see Sect. 2.2.4).

Once all spaxels have their wavelength-spectrum, we merged them into a unique cube using the nightly HexaGrid for their $\{x, y\}$ position into the focal plane; x increasing northwards and y increasing westwards.

2.2.4. Wavelength flexure correction

Once a cube has been created, we can test and correct the wavelength solution accuracy by comparing the measured sky and telluric line positions with their expected locations. We first randomly arranged the 600 faintest spaxels (≈half of the MLA) into 40 spectra averaging 15 spaxels. We assumed the sky to be constant over the IFU given its small field of view. The faintest spectra are thus the ones for which the sky ratio is the largest. The random spaxel mixing enables us to average out any potential spatial structure. This way all 40 spectra are comparable.

We fit these 40 spectra with the sum of a fifth-degree polynomial and two Gaussian lines: one in emission for the strong sky sodium line, and one in absorption, for the main O_2 telluric complex. The centroid of these lines is allowed to shift. We show in Fig. 7 the difference between the expected measured lines and their positions, assumed to be 5892.3 Å and 7624.5 Å for the sodium and the telluric lines, respectively. We considered the effective wavelength shift to be the average shifts, weighted by the square of the inverse of the mean shift error. The wavelength shift, $\Delta\lambda$, is then converted into a horizontal pixel shift, Δi. In the example shown in Fig. 7, this shift is as small as +0.01 pixels. Typical Δi shifts are ≈0.5 pixels, and are rarely above one pixel.

We note from Fig. 7 that the two line shifts are not identical. While this is not always the case, it actually is expected. The O_2 telluric absorption is a complex combination of numerous blended lines whose relative amplitudes depend on atmospheric conditions (see Sect. 3.3). Therefore the exact telluric line centroid varies, which makes it a questionable reference candidate. We nonetheless decided to use it for two reasons: First and most

importantly, the telluric O_2 complex is always clearly detected, even during short exposures and for cases in which the sky brightness is small in comparison to the targeted source – including its host environment. Consequently, keeping in mind that PYSEDM is designed to be a fully automated pipeline and given the low wavelength resolution of the SEDM, the reliability of the detection of the telluric line largely counterbalances any inaccuracy of a few angstroms. Second, because we are applying a single global pixel shift as a wavelength flexure correction, using the telluric line along with the sodium sky line enables us to have two wavelengths to anchor the wavelength shift. We note that part of the discrepancy observed in Fig. 7 could also be due to the wavelength flexure correction being more complex than the simple pixel shift assumed in PYSEDM. However, given the low wavelength resolution of SEDM, this instrument should not be used for accurate redshifts or element-velocity analyses and we therefore decided to keep the simple and reliable single shift technique.

3. Spectral extraction from 3D cube

Within two minutes of the end of a science exposure, a 3D cube fits file is created following the PYSEDM cube extraction detailed in Sect. 2. Each spaxel of the 3D data cube includes three intensity components, each one characterized by its own spatial signature: (1) the target flux, from a point source located close to the center of the MLA, (2) the night sky flux, a spatially flat component over the entire IFU field of view, and (3) the (potential) host galaxy flux, a structured background. The core purpose of PYSEDM is to automatically extract the point source spectrum to enable the typing of the transient with no human interaction. However, for full flexibility and reliability, the code allows human interaction for potential fine-tuned re-extractions of complex cases (see Sect. 3.5).

In the following sections we present the transient spectrum-extraction process. PYSEDM represents an improvement in background subtraction over the pipeline that was presented in Sect. 4.2.3 of Blagorodnova et al. (2018). With PYSEDM we are able to automatically extract the target spectrum from a single SEDM exposure, unlike the previous pipeline which required an A/B offset pair to subtract the background accurately. Removing the overhead needed to acquire an A/B offset pair substantially improves the observing efficiency of SEDM. We start Sect. 3.1 by showing how we made use of the imaging channel of the SEDM to estimate the target location in the MLA. Then, in Sect. 3.2 we describe the 3D point spread function (3D-PSF) modeling and the actual source extraction. The flux calibration procedure is described in Sect. 3.3, and in Sect. 3.5 we describe the manual operation of the transient extraction.

The PYSEDM point source extraction pipeline is largely inspired by the Nearby Supernova Factory (Aldering et al. 2002) collaboration development of the SuperNova Integral Field Instrument (SNIFS, Lantz et al. 2004); see also the detailed work presented (in French) in Copin (2013). We have built a simplified version of their 3D-PSF extraction (see Sect. 3.2.1). We have also adapted their telluric absorption correction method (see Sect. 3.3, and Buton et al. 2013), and we are using their atmospheric differential refraction correction algorithm (see Sect. 3.2.3, and Copin 2013).

3.1. Target location in the multi-lens array

– Input: [guider_files_of_crr_target.fits];
– Output: metaguider_astrometry_crr_target.fits.

The P60 telescope pointing is not accurate enough to perfectly place the target in a pre-defined location on the MLA. Furthermore, we cannot directly solve the astrometric solution of the MLA, as it rarely contains more than the target and its potential host in the 28×28 arcsec^2 field of view. It is therefore non-trivial to know where the target is located without inspecting the data cube – which would make the pipeline non-automatic.

The simplest cases are those in which the target is the brightest or the sole visible object in the MLA. Observations of standard stars are a typical example. For them, one can simply take the brightest spaxel as a good initial guess of the target location at a reference wavelength (see Sect. 3.2.3). The spaxel brightness is defined as the integrated flux between two wavelength boundaries – typically 5000–8000 Å for PYSEDM. In practice, to ensure the robustness of the automatic pipeline, we have not used the brightest spaxel position, but the median position of the five brightest. This avoids single-spaxel issues, such as cosmic ray contamination. This simple target positioning is referred to as the "brightness position" method.

General cases are the most complex. A typical target is accompanied by and possibly overlapped by a host galaxy detectable in the MLA. In many cases the targeted transient is not the brightest object in the MLA. For such cases, we need external information to automatically locate the target position. Fortunately, the SEDM has a parallel imaging channel. This channel is used for guiding during IFU exposures. As detailed in Blagorodnova et al. (2018), this photometric channel, called the rainbow camera (RCAM), has a 13×13 arcmin^2 field of view split into four quadrants, each having a different optical filter (u, g, r, i) and having the IFU pick-off mirror mounted in the center. Thus, the RCAM is illuminated from the P60 focal plane simultaneously with the 28×28 arcsec^2 field of view of the IFU. The position of the IFU pick-off prism is fixed, allowing one to project a WCS solution calculated from the guider images onto the MLA. This is what is done in PYSEDM.

At the end of each science acquisition, we created a median stack of all guider images (one taken every 30 s), and we measure the astrometric solution of this "meta-guider" image using a local installation of the astrometry.net software (Lang et al. 2010). Since we know the target coordinates, we are able to predict the RCAM CCD pixel coordinates $\{x_{rc}, y_{rc}\}$ where the target should be located. We note that, if the telescope pointing is correct, the target will not be visible in the RCAM image since its light has been reflected by the pick-off mirror onto the MLA. Thus, we used the following transformation to project RCAM pixel coordinates $\{x_{rc}, y_{rc}\}$ onto MLA cube coordinates at a reference wavelength of 7000 Å $\{x, y\}$:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} p_1 & p_2 \\ p_3 & p_4 \end{bmatrix} \begin{bmatrix} x_{rc} - x_0 \\ y_{rc} - y_0 \end{bmatrix}, \tag{1}$$

where p_i, the transformation parameters, and $\{x_0, y_0\}$, the centroid shift coordinates, have been derived based on observations of tens of standard stars for which $\{x, y\}$ coordinates have been measured using the "brightness position" method. The IFU pick-off mirror is only 5 mm square and is rigidly mounted to the central support of the RCAM filter assembly, so we expect no significant flexure in its position relative to the RCAM images. Therefore these six parameters are fixed but nonetheless must be updated if any part of the optical system (including the CCDs) is adjusted.

With this method, referred to as the "guider-astrometry position" method, we are able to predict the position of the transient on the MLA to within a couple of arc seconds. This is good

enough to avoid confusion when automatically extracting a point source within its environment (see Sect. 3.2.1). The "guider-astrometry position" method is illustrated in Fig. 5, where we see that the expected point source position, marked as an "x", appears about 1 spaxel (≈ 0.75 arcsec) southwest from the true PSF centroid.

To ensure the PYSEDM pipeline robustness, if the "guider-astrometry position" method fails for any reason – could not access the guider images, astrometric solution of the meta-guider did not converge, etc. – we then used the "brightness position" method as a backup solution. In such cases – less then 5% of the time – a warning is placed in the header as well as on the validation plots (see Sect. 3.6). These cases usually require manual re-extraction if the transient is faint (see Sect. 3.5). On rarer occasions, the P60 acquisition offset fails, and we find with the "guider-astrometry position" method that the target is outside the MLA. In this case a warning is raised, the target spectral extraction outputs a null spectrum, and PYSEDM reports the target as "failed". Most of the time, this is caused by poor weather conditions.

3.2. 3D PSF extraction

- Input:
  - e3d_target_file.fits;
  - "source position" (automatic or manual);
- Output:
  - spec_target_file.fits;
  - e3d_psfmodel_file.fits.

The 2D PSF profile modeling is presented in Sect. 3.2.1. We then detail in Sect. 3.2.2 how this profile is actually matched to the data accounting for PSF ellipticity and target background. In Sect. 3.2.3, we explain how we modeled the evolution of the PSF centroid as a function of wavelength. The complete spectral extraction procedure is detailed in Sect. 3.2.4.

3.2.1. PSF profile

Given an IFU wavelength-slice or meta-slice, and calling r the elliptical radial distance from the PSF center, we parameterize our radial PSF function $\mathcal{P}$ as a linear combination of a Gaussian and a Moffat distribution:

$$\mathcal{P}(r; \alpha, \sigma, \rho_{nm}) = \frac{\rho_{nm}}{1 + \rho_{nm}} \times \mathcal{N}(r; \sigma) + \frac{1}{1 + \rho_{nm}} \times \mathcal{M}(r; \alpha, \beta(\alpha)), \tag{2}$$

where $\mathcal{N}(r, \sigma)$ is the normalized Gaussian distribution center of 0 with a scale of σ – from scipy.stats.norm – and $\mathcal{M}(r, \alpha, \beta(\alpha))$ is the normalized Moffat distribution expressed below; ρ_{nm} is the relative amplitude of the two distributions:

$$\mathcal{M}(r; \alpha, \beta) = 2 \frac{\beta - 1}{\alpha^2} \left[1 + \left(\frac{r}{\alpha} \right)^2 \right]^{-\beta}. \tag{3}$$

An illustration of the "Gaussian+Moffat" PSF model is presented in Fig. 8; see details of the complete model in Sect. 3.2.2.

To limit the number of correlated parameters in the PSF determination, the Moffat β parameter in Eq. (2) is parameterized as a linear function of α, such that $\beta = b_0 + \alpha \times b_1$ as inspired by Buton et al. (2013). b_0 and b_1 have been determined from high signal-to-noise ratio standard star observations to be 0.25 and 0.63, respectively. These values are in agreement with the corresponding parametrization for SNIFS presented in Buton et al. (2013).
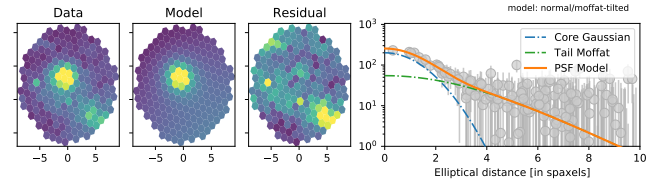
Fig. 8. Illustration of the PSF model used to extract to target point source from the cubes 3D, here showing the slice [5000, 5500] Å of the part of ZTF18abqlpgq's x, y, λ-cube used for PSF fitting (see text). *From left to right*: data slice, model slice (PSF+tilted-plane background), residual slice and PSF profile. The PSF profile axes on the right hand side present the background-subtracted spaxel flux (gray marker with error bars) [in pseudo-ADU, see text] as a function of the elliptical distance to the best fitted center of the PSF. The orange curve is the best fitted PSF profile, decomposed into Moffat (green) and Gaussian (blue). The small structure visible on the lower right of the residual slice panel is the host galaxy signal (see Fig. 5).

3.2.2. PSF+background slice model

Our primary aim is rapid typing of the observed transient with the SEDM. Thus, we are not taking final reference images, in other words, observation of the host galaxy long after the transient has faded away. Therefore, we cannot apply the host removal technique presented in Bongard et al. (2011).

Instead, to account for the spatially flat (sky) and structured (host) background signals associated with the target point source, we fit a 2D multi-degree polynomial simultaneously with the PSF profile, as done in Bailey et al. (2009). Three levels of complexity have been implemented: a simple flat continuum, a tilted-plane, and a curved-plane. These have one, three and five free parameters, respectively. The default background in PYSEDM is the tilted plane (Bailey et al. 2009) and will be the one used in this paper if not specified otherwise.

Altogether, the "Gaussian/Moffat + Tilted Plane" 2D model of a cube wavelength-slice (or meta-slice) has 11 (± 2) free parameters:

- Two for the PSF centroid position $\{x_{psf}^0, y_{psf}^0\}$;
- Two for e and θ, the PSF ellipticity and ellipse rotation parameter, respectively;
- One for the PSF profile amplitude (the flux);
- Three for the normalized PSF radial profile (see Sect. 3.2.1);
- Three for the background (1 if flat continuum, 5 if curved plane).

Using this model, we are able to create a 2D slice model by assuming that each slice "spaxel" has a flux value corresponding to that of its center in $\{x, y\}$ coordinates. For the radial parametrization of the PSF, we convert $\{x, y\}$ to the centroid elliptical coordinate defined by the radial distance $r = \sqrt{x_{ell}^2 + y_{ell}^2}$ which gives the following transformation to x_{ell} and y_{ell}

$$\begin{bmatrix} x_{ell} \\ y_{ell} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x - x_{psf}^0 \\ y - y_{psf}^0 \end{bmatrix}. \tag{4}$$

An illustration of the "PSF+Tilted Plane" model is presented in Fig. 8. We see in the right panel of the figure that the "Gaussian+Moffat" profile well describes the data: the Gaussian profile dominates the core of distribution while the Moffat profile well matches the PSF tail. We note that a good PSF radial match also implies a good fit of both the PSF centroid $\{x_{psf}^0, y_{psf}^0\}$ and of the PSF elliptical parameters e and θ.

The "Model" panel represents the model-slice reconstruction where one can distinguish the 2D reconstruction of the PSF on

top of the tilted plane. We note the southwest to northeast inclination of the background. The goodness of fit is demonstrated in the "Residual" panel where we see no structure left at the PSF location. The bright signal in the lower-right (also visible in the "Data" panel) is actually a part of the host galaxy (see Fig. 5). As further discussed in Sect. 3.2.4, only the part of the cube near the target is fitted.

3.2.3. ADR correction

The atmosphere acts as a prism and bends light by an amount that depends on wavelength. The image is then shifted from its original position $\{x_{ref}, y_{ref}\}$ toward the zenith to a degree dependent on the wavelength, the atmosphere refraction index n, and the target airmass z. To first order, the mean atmospheric refraction is accounted for by the telescope pointing model. Hence, only the atmospheric differential refraction (ADR), related to wavelength variations of the refraction effect, has to be taken into account during the spectral extraction from the 3D cube: see details in Copin (2013).

Theoretically, three variations affect the observations: (1) temporal variations – the evolution of n and z as a function of time during the exposure, (2) spatial variations – the variations of n and z within the field of view, and (3) chromatic variations – the wavelength dependence of n. In the SEDM, the spatial variations are negligible. With less than one-hour exposures, the temporal variation will induce a slight blurring that will be, to first order, accounted for by the PSF model. Consequently, only the chromatic variation has to be modeled.

The effect of the chromatic ADR has been extensively studied (e.g., Filippenko 1982), including specifically for 3D spectroscopy (e.g., Arribas et al. 1999). The target position in the MLA as a function of wavelength $\{x(\lambda), y(\lambda)\}$ is thus given by

$$x(\lambda) = x_{ref} - \frac{1}{2}\left(\frac{1}{n^2(\lambda)} - \frac{1}{n^2(\lambda_{ref})}\right) \times \tan(d_z)\sin(\theta), \quad (5)$$

$$y(\lambda) = y_{ref} - \frac{1}{2}\left(\frac{1}{n^2(\lambda)} - \frac{1}{n^2(\lambda_{ref})}\right) \times \tan(d_z)\cos(\theta), \quad (6)$$

where $d_z = \arccos(z^{-1})$ is the zenith distance, θ is the parallactic angle, and λ_{ref} is a reference wavelength set at 7000 Å in PYSEDM. Hence, $\{x_{ref}, y_{ref}\}$ is the target position in the MLA at λ_{ref}. We use the air refraction index $n(\lambda)$ modeled by Stone & Zimmerman (2011)[4], which depends on pressure, temperature and relative humidity. These atmospheric parameters are provided in the FITS image header.

Deriving the $\{x(\lambda), y(\lambda)\}$ consequently requires four free parameters: the reference position $\{x_{ref}$ and $y_{ref}\}$, the effective airmass, z of the exposure, and the parallactic angle, θ. The ADR fitting procedure is detailed in Sect. 3.2.4.

3.2.4. Point source spectral extraction

The point source extraction is made in three steps: (1) fit the "PSF+Background" model on several large wavelength meta-slices, (2) model the run of the PSF parameters with wavelength, and (3) fit the PSF amplitude (together with the background) at each wavelength slice of the 3D cube. The derived PSF amplitude as a function of wavelength is equivalent to the target spectrum.
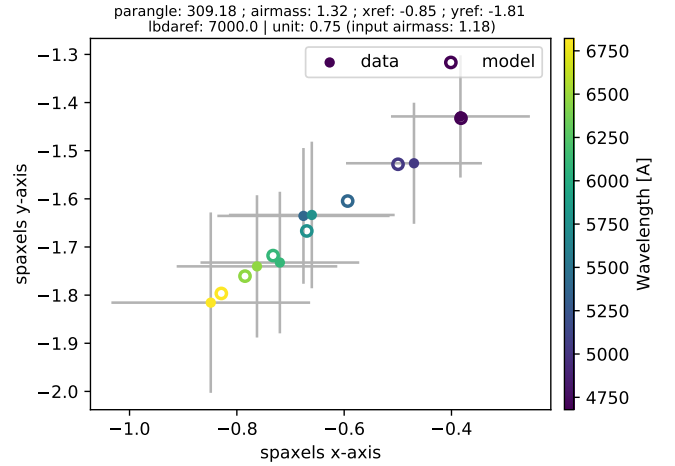
---

4 NIST: https://emtoolbox.nist.gov/Wavelength/Documentation.asp

Fig. 9. Best-fit PSF centroid positions (full-markers) estimated on six meta-slices of ZTF18abqlpgq's x, y, λ-cube during the first stage of the automated PSF extraction. Marker colors indicate the effective wavelength of the slice. Atmospheric differential refraction causes the true position of the target in the MLA to vary with wavelength. Our best model of this evolution is shown as open markers. This automated ADR parameter estimation allows us to predict the target centroid to an accuracy of a few percent of a spaxel.

For step (1), we split the 3D cube into six meta-slices with regular wavelength bins between 4500 Å and 7000 Å. This wavelength range is where the SEDM's efficiency is at its maximum (Blagorodnova et al. 2018), yet extends blue enough to well constrain the ADR parameters. We note that the number of meta-slices and the wavelength range used can be manually adjusted by the user in the case of a non-automatic extraction. The rightmost panel of Fig. 8 shows the PSF profile fit of the second bluest meta-slice (4916, 5333 Å) made during the automatic extraction of the ZTF18abqlpgq spectrum.

For step (2), we fit each of the PSF parameters as a function of wavelength using the meta-slices from the first step. This contains two parts: the ADR component and the PSF profile component. For the ADR component, we use the PSF centroids $\{x_{psf}^\lambda, y_{psf}^\lambda\}$ measured per meta-slice (made during step (1)) to fit the ADR free parameters $\{x_{ref}, y_{ref}\}$, z, and θ (see Sect. 3.2.3). During the fit, z and θ are bounded by their corresponding measurements stored in the header at the beginning and the end of the exposure. The reference position, $\{x_{ref}, y_{ref}\}$, should be within three spaxels of the expected target position (see Sect. 3.1). We use the best fitted ADR parameters in the next step to fix the PSF centroid position as a function of wavelength. As illustrated in Fig. 9 we were able to accurately predict the PSF centroid position as a function of wavelength to better than a tenth of a spaxel. This is further evidence that the PSF model used is reasonable, as the physical modeling of the ADR matches the observed centroid evolution. For the PSF profile component, we use the measured profile values for each meta slice to model their behavior as a function of wavelength. The ellipticity parameters e and θ are assumed to be fixed with wavelength and are therefore set to their mean values. To model the dependence of the remaining three parameters α, σ, and ρ_{nm} on wavelength, we use standard stars which are bright enough to divide into many meta-slices and still obtain a good PSF fit. The Moffat α and the ratio parameter ρ_{nm} turn out to be largely achromatic. While a linear fit would have been more flexible, we decided to use the achromatic approximation to increase the PYSEDM automatic pipeline robustness.
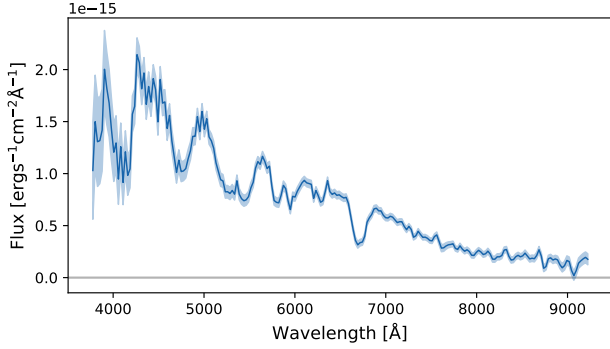
Fig. 10. Flux-calibrated spectrum of ZTF18abqlpgq automatically extracted using PYSEDM. The shaded blue band represents the flux error.

The Gaussian σ parameter is evolving as $\sigma(\lambda) \simeq \sigma_{\text{ref}} (\lambda/\lambda_{\text{ref}})^{-1/5}$ as expected from Kolmogorov seeing modeling. This is expected, since the Gaussian traces the core of the PSF profile (Fried 1966; Tokovinin 2002).

In step (3), we perform the actual flux extraction. The chromaticity of the PSF parameters, including centroids and ellipticity, were fixed in step (2). The sole remaining free parameters are the background parameters and the PSF intensity. These are fitted iteratively for each wavelength slice. The PSF amplitude as a function of wavelength corresponds to the target spectrum in the cube's native, that is, non-flux-calibrated, units.

The resulting 3D PSF model cube and the target spectrum are saved as fits files after flux calibration (see Sect. 3.3). The extracted flux-calibrated spectrum of ZTF18abqlpgq is shown in Fig. 10. This r-band 19 mag spectrum is a typical good example of what "SEDM+PYSEDM" can automatically acquire in 2430 s.

3.3. Flux calibration

– Input files: calspec_file;
– Input pipeline product: spec_target_file.fits;
– Created product: fluxcalibration_target_file.fits.

The spectrum extracted at the end of Sect. 3.2.4 are in units we refer to as "pseudo" ADU: "pseudo", because several renormalizations have been applied during the extraction process. The purpose of the flux calibration step is to recover the target spectrum in physical units. This means removing the instrumental response $C(\lambda, t)$ and the telluric absorption $\mathcal{T}_{\text{atm}}(\lambda, t, z)$, which depends on airmass z. We developed a much-simplified version of the method from Buton et al. (2013), since our goal is focused on rapid and reliable transient classification rather than accurate spectrophotometry.

In the late twilight, as well as a few times during the night (varying depending on the schedule), we take short observations of standard stars from the calspec catalog (Bohlin et al. 2014). The flux calibration step of PYSEDM compares calspec-calibrated spectrophotometry to the observed SEDM spectra to derive the instrumental response of the SEDM and to solve for the telluric correction for the standard observation.

During the automatic pipeline processing, each science target is flux-calibrated using the C from the telluric-corrected standard star observation and is also telluric-corrected using $\mathcal{T}$ based on its own airmass.

Following Buton et al. (2013), we use the high resolution telluric spectrum from the Kitt Peak National Observatory[5]. This spectrum is split into two categories: the O_2 wavelengths, and
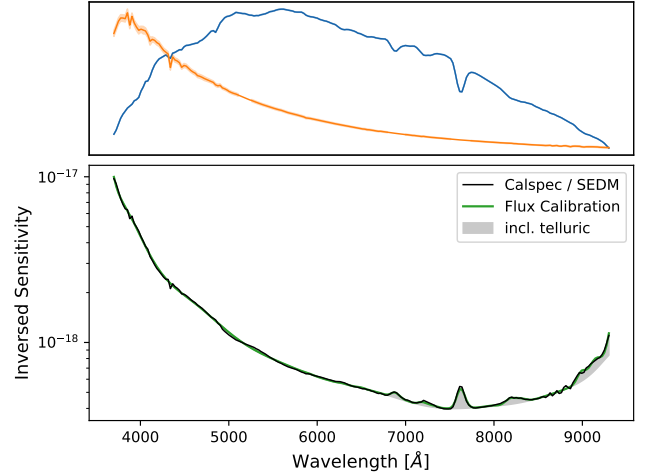
Fig. 11. Illustration of the flux calibration procedure. *Top*: in blue, extracted standard star spectrum in pseudo-ADU of BD+33d2642; in orange, flux-calibrated CalSpec spectrum of the same star convolved at the SEDM wavelength resolution used as flux reference. *Bottom*: in black, ratio between CalSpec and SEDM spectra (top). The green curve is the full flux calibration model (continuum+telluric). The influence of the telluric correction is shown as a gray shaded area; hence the lower of the gray area represents the continuum component of the flux calibration. When flux-calibrating a science target, the intensity of the telluric component (in gray) varies to match that expected at the science target air-mass.

the H_2O wavelengths (see Table 1 of Buton et al. 2013). Wavelengths not contained in either of the two components are assumed to be free from telluric absorption. The relative amplitude c_i and the airmass dependence ρ_i are free parameters for each of the two parts ($i = O_2$ or H_2O), such that

$$\mathcal{T}(z) = \mathcal{T}_{O_2} \times (c_{O_2} \times z^{\rho_{O_2}}) + \mathcal{T}_{H_2O} \times (c_{H_2O} \times z^{\rho_{H_2O}}), \qquad (7)$$

where z is the airmass and $\mathcal{T}_i$ is the telluric spectrum for which the fluxes of wavelengths not contained within the corresponding wavelength-ranges are set to 0. We model C as a 20th degree Legendre polynomial.

The estimation of C and $\mathcal{T}$ parameters is made as follows. Given a standard star observation S in pseudo-ADU, we load its corresponding calspec spectrophotometric spectrum that we convolve to the SEDM resolution S_{Ref}. We then simultaneously fit for the instrumental response C and telluric $\mathcal{T}$ parameters by minimizing the difference between the S/S_{ref} ratio and the inverse sensitivity curve defined as the sum $C + \mathcal{T}$. The best fitted inverse sensitivity curve of an observation of the calspec standard star BD+33d2642 is shown in Fig. 11.

We store the four telluric parameters and the C inverse sensitivity spectrum, assumed to be airmass-independent, as a FITS file. When flux-calibrating a science target spectrum, we used the target airmass rather than the standard star airmass in Eq. (7).

To test the quality of the flux calibration, we examine the ratio between the reference spectrum and the flux-calibrated observed spectrum for any standard star. In Fig. 12, we present the ratio between SEDM flux-calibrated standard star spectra and their corresponding calspec spectra convolved to match the SEDM wavelength scale. The flux ratios encapsulating 0.95–1.05 (pale blue) and 0.84–1.16 (light gray) are indicated as well as the flux ratio median (dark blue) for the 141 standard star observations taken in October 2018 used for this test. The plot shows the rms scatter of the median ratio is around 1%, while the range of the 1σ scatter stays within an envelope of a few
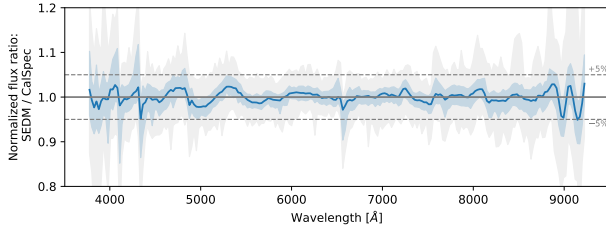
Fig. 12. Mean-normalized ratio between flux-calibrated standard star spectra (considered as science targets, various stars have been used) and their expected CalSpec spectra. The solid blue line shows the running median spectrum value, while the blue and gray bands encapsulate 68% and 95% of the spectra, respectively. This figure illustrates that our color calibration is accurate at the few percent level.

percent. It is therefore expected that our flux-calibrated science spectra are relatively flux-calibrated at the few-percent level. A more thorough analysis would be necessary for the spectrophotometric absolute calibration, but this is both beyond the scope of this paper and the SEDM purpose in general. More generic tests of the pipeline results will be presented in Sect. 4.

3.4. Accuracy of the wavelength calibration

In Sect. 2.1.2 we presented the construction of the nightly wavelength solution that is then fine tuned for each science exposure to account for flexure; see Sect. 2.2.4. In this section, we compare PYSEDM extraction of HZ standard star spectra with their corresponding CalSpec spectra. The upper panel of Fig. 13 shows an SEDM spectrum overplotted with a CalSpec spectrum smoothed to match the SEDM wavelength resolution. As shown in this figure, HZ stars have strong hydrogen absorption lines, which enable us to test the wavelength accuracy of the PYSEDM pipeline. To do so, we fit the absorption line region [3900, 5000 Å] for a wavelength offset between the CalSpec spectrum and the SEDM spectrum together with a 2D (linear) multiplicative polynomial to account for potential flux calibration inaccuracy. We repeated this procedure for the 94 HZ4 star observed over the one-month period in September, 2018. The lower panel of Fig. 13 shows the histogram of the fitted wavelength shifts. The rms of this distribution is ~3 Å, or ~0.1 SEDM pixels. For a typical wavelength of 6000 Å this corresponds to a precision of 4×10^{-4} in redshift, or 120 km s^{-1}. Despite this high precision of the PYSEDM pipeline, Fig. 13 shows a systematic offset of ~3 Å toward the blue. The origin and correction for this small, but significant, offset is under investigation.

3.5. Manual extraction option

The pipeline is fully automated and needs no manual input from the data acquisition to the generation of a flux-calibrated spectrum. However, one can manually call any of the PYSEDM functions to bypass some of the automated steps. In particular, the target position and the part of the SEDM cube used to run the 3D PSF extraction can be specified on the command line by simply providing {x, y} (the target reference centroid position guess otherwise given by the astrometric procedure detailed in Sect. 3.1). An interactive plot frontend also exists to manually click on the expected target location and draw the extraction region. Further details on how to interact with PYSEDM is given in the online documentation[6].
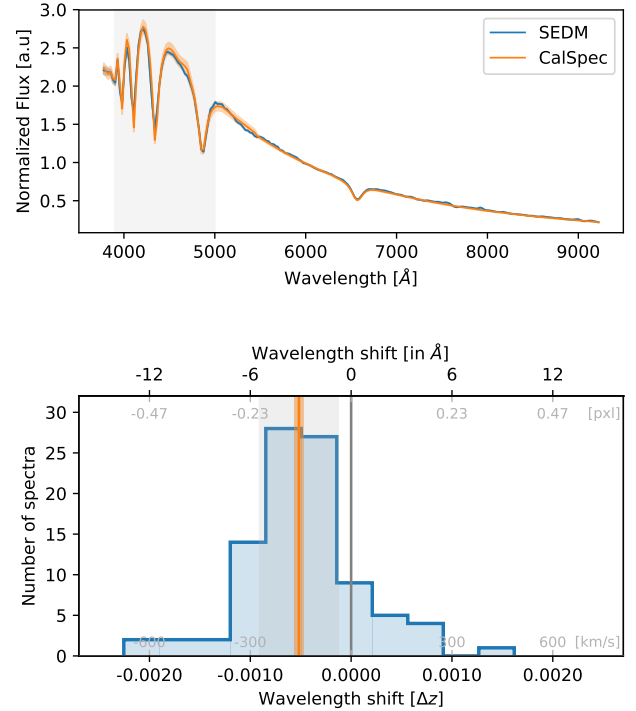
Fig. 13. Accuracy of the wavelength calibration. *Top*: SEDM and a CalSpec spectrum of the HZ4 standard stars. The gray band shows the wavelength region used to fit for the wavelength shift between the two spectra. *Bottom*: histogram of the wavelength shift fitted on the 94 HZ4 star spectra obtained by SEDM in September 2018. The lower axis presents the shift in redshift shift (in z or in km s^{-1}), the upper axis shows the corresponding offset in Å or pixel for a typical wavelength of 6000 Å. The vertical orange line represents the mean and error on the mean. The light-gray band shows the standard deviation. SEDM spectra automatically extracted by PYSEDM have a precision of ~3 Å or 4×10^{-4} in redshift. A small 1 s.d. offset is visible suggesting that the SEDM spectra a lightly blue-shifted of approximately a tens of a pixel. The origin of this shift is under investigation.

3.6. Quality flag, typing and check plots

At the end of the automated processing, we apply the following criteria to generate a quality flag value that is recorded in the output header with the keyword QUALITY:
– 5: the automated target positioning in the MLA (see Sect. 3.1) failed and the brightest object in the field was extracted by default,
– 4: low object signal (defined as more than 20% of the flux-calibrated spectrum is negative),
– 3: the automatic target positioning placed the target outside the MLA,
– 2: not assigned yet,
– 1: A quality 5 spectrum that was determined to be correctly extracted by human inspection of the field and the SNID fit and then manually recovered,
– 0: no problem.

If the science target is not a standard star, and if its quality flag is 0, we run SNID (Blondin & Tonry 2007) on the flux-calibrated spectrum to make an automated classification. If the SNID "rlap" of the first match is greater than 4.0 we push this classification. Automatic classifications should be used with care.

At the end of each exposure, a summary plot encapsulating all the relevant automatic extraction plots is saved and pushed to the SEDM clients, including the ZTF Growth Marshal. We
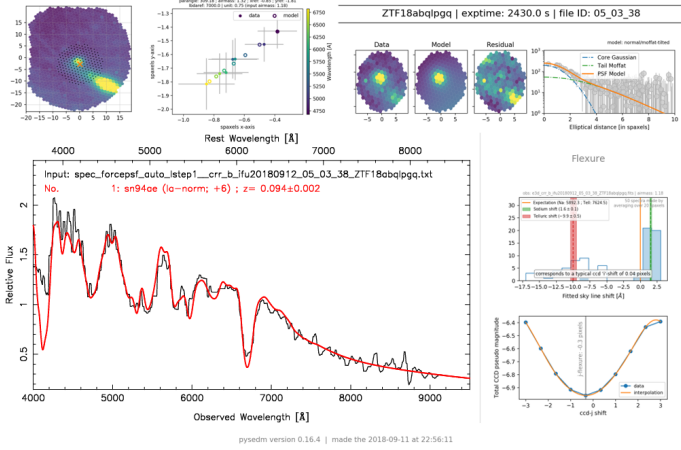
Fig. 14. Summary figure presenting in one compact image the relevant plot automatically generated by the PYSEDM pipeline. The figure is automatically created and pushed together with the final flux-calibrated spectrum to registered clients, notably the ZTF marshal. This example is a science observation of ZTF18abqlpgq.

show in Fig. 14 the "pysedm-report" automatically created for the observation of our example Supernova ZTF18abqlpgq. The SNID match is excellent with a Type Ia supernova at six days post-maximum at a redshift of $z \sim 0.094$.

4. Results

4.1. Some statistics for 19 weeks of operations

Since the beginning of August 2018, PYSEDM has been the sole pipeline for the SEDM. The ZTF collaboration is using this instrument to type the transients discovered by the ZTF camera mounted on the Palomar 48″ Schmidt (P48) telescope. We present in Fig. 15 the fraction of object types generated by PYSEDM as a function of time, and the fraction of time spent acquiring SEDM observations. We spend, on average, 6h30m on targets each night. In addition, we have a typical overhead of 180 s, including slew time and readout. We point out two dips in the observing time; the deepest in mid-October and a smaller dip in early December. These correspond to an engineering period (mid-October), and a period of poor weather conditions, notably due to wildfires in southern California, at the end of November to early December 2018.

About two thirds of the SEDM observing time is spent acquiring supernova spectra, two thirds of which are Type Ia supernovae. We see in Fig. 15 that the fraction of time spent on standard stars (90% of "other incl. std") has continuously been reduced from ~15% to about 2%. The reasons are threefold: first, we observed numerous standard stars when launching PYSEDM in production to asses the robustness of the pipeline; second, at that time not all ZTF fields had references, which limited the amount of SEDM triggering for transients; and third, the human transient scanning has progressed optimizing the triggering of extra-galactic candidates. We also note that the fraction of variable stars ("CV") is reducing as a function of time. Except for rare occasions, such targets are purposefully not observed, hence the time spent observing them is reducing as ZTF triggering of the SEDM machine is improving. Finally, one can notice that the fraction of time spent acquiring spectra that turn out to be classified as "bad" (quality $\neq 0$, see Sect. 3.6) is slightly increasing with time. This is due to observing conditions getting worse while entering the winter period.
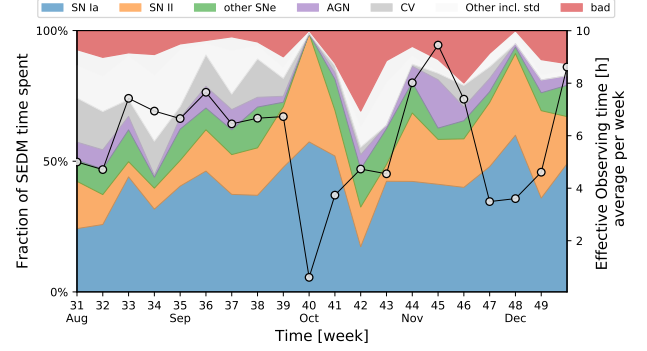
Fig. 15. Use of the SEDM time since PYSEDM is the sole SEDM pipeline (early august 2018) and until redaction of this paper (19 weeks). Color bands show the fraction of target observed per week sorted by type (see legend): left axis. The running line with gray markers shows the evolution of the effective time spent on target per night (weekly average): right axis.
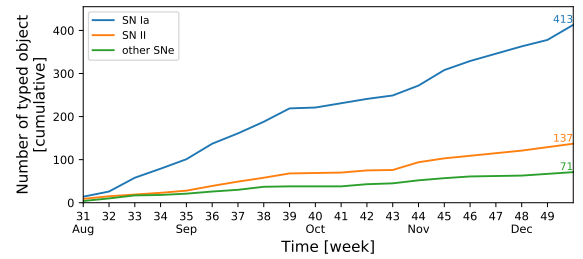

Fig. 16. Number of supernovae typed by the SEDM as a function of time.

Figure 16 presents the number of individual supernovae typed by the SEDM since the launch of PYSEDM in early August. In 19 weeks, we have typed ~400 SNe Ia, almost ~140 SNe II and ~70 other SNe, mainly Ib/c. This corresponds to roughly 20 new SNe Ia and 7 new Type II per week. See classification details and more accurate discovery estimation in Fremling et al. (in prep.).

4.2. Spectral accuracy: Comparison with other spectrographs

In some (rare) instances, a target observed by the SEDM has also been observed by another spectroscopic facility at close to the same time, typically with higher spectroscopic resolution. We used this opportunity to test the quality of the PYSEDM spectral extraction. In Fig. 17, we show four representative spectral comparisons for each of three supernova types: Type Ia, Type II and Type Ib/c. The objects were selected to be representative of our sample spanning different phases, redshifts and magnitudes. The characteristics of these spectra are summarized in Table 1.

This figure demonstrates the quality of the SEDM spectra extracted using PYSEDM. For each case, the overall shape matches that of the other spectrum or spectra, which confirms the accuracy of the color calibration as discussed in Sect. 3.3. In addition, the main characteristic features, such as the silicon line at 6500 Å for the SNe Ia, or the P-Cygni Hα shape for the Type II, are clearly visible in the SEDM spectra. This shows that a very low-resolution spectrograph is well suited for transient typing (Blagorodnova et al. 2018). However, galaxy emission lines, as seen in the spectra of ZTF18abukavn, are not well resolved by the SEDM (see details concerning this target in Ho et al. in prep.). This could limit our ability to accurately derive host redshifts from SEDM IFU data as done, for example, with
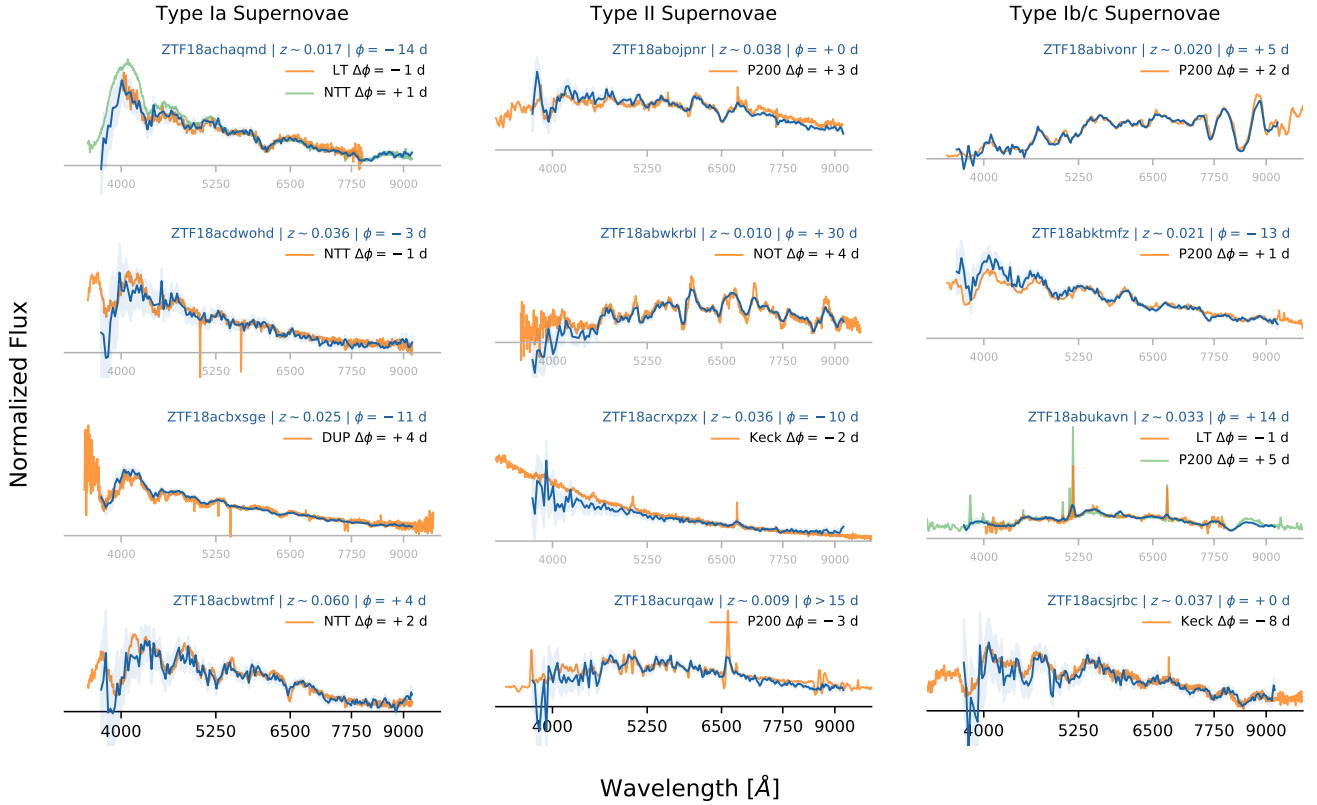
Type Ia Supernovae   Type II Supernovae   Type Ib/c Supernovae

Fig. 17. Comparison between SEDM spectra (in blue) with other higher resolution instruments. Legends in each panel present: in blue the target name, redshift and phase of the SEDM spectrum; in color (orange and/or green), the other facility name and the phase difference with respect to the SEDM spectrum. For readibility, the *x*-axis of each sub-panel is aligned with the zero of the corresponding *y*-axis. Facility names: NTT (New Technology Telescope, ePESSTO, see Smartt et al. 2015, http://www.pessto.org), LT (Liverpool Telescope JMU; SPRAT Piascik et al. 2014), DUP (LCO-duPont ; WFCCD), P200 (Palomar 200-inch Telescope; Double Spectrograph), Keck (Keck Telescope; LRIS), NOT (Nordic Optical Telescope; Alhambra Faint Object Spectrograph and Camera). For visibility P200 spectra have been smoothed by a five-bin Gaussian filter.

Table 1. Summary of supernovae presented in Fig. 17.

ZTF name	TNS name	RA (deg)	Dec (deg)	Redshift	Typing	Phase (days)	<Magnitude ("ztf:g")	Exptime (s)
ZTF18acbwtmf	SN2018hrn	322.8909	+23.0147	0.060	Ia	4	18.1	1200
ZTF18acbxsge	SN2018huw	9.8009	−5.2007	0.025	Ia	−11	17.6	1600
ZTF18acdwohd	SN2018ids	319.0474	+11.9462	0.036	Ia	−3	16.5	1600
ZTF18achaqmd	SN2018ilu	353.3374	+4.8096	0.017	Ia	−14	17.2	1200
ZTF18acurqaw	SN2018hwm	125.3675	+3.1645	0.009	II	>15	19.0	2250
ZTF18acrxpzx	SN2018jrj	167.7667	+64.2474	0.036	II	−10	18.3	2250
ZTF18abwkrbl	SN2018gjx	34.0649	+28.5913	0.010	IIb	30	18.5	1440
ZTF18abojpnr	SN2018fzn	297.4871	+59.5928	0.037	IIb	2	19.3	2430
ZTF18acsjrbc	SN2018jme	96.1574	+50.9750	0.037	Ic	0	19.1	2250
ZTF18abukavn	SN2018gep	250.9509	+41.0454	0.033	Ic-BL	12	17.9	1440
ZTF18abktmfz	SN2018eoe	336.1015	+18.0719	0.020	Ib	−13	18.8	2430
ZTF18abivonr	–	321.8796	+45.4223	0.020	Ic	5	19.4	2430

Notes. Phases and magnitudes are approximations derived from light curve fitting using sncosmo. They are illustrative and should be used with care, especially for Type II supernovae.

SNIFS (Rigault et al. 2013, 2018); investigation into improvement is ongoing.

5. Summary and conclusions

This paper presents PYSEDM: a fully automated pipeline for the SEDMachine (Blagorodnova et al. 2018). First, we detailed the procedure for constructing an x, y, λ 3D-cube from the 2D CCD exposures acquired by the integral field unit. We presented the construction of the wavelength solution, the trace identification, the flat fielding, and the flexure corrections. Then, we detailed the automatic target point source spectral extraction from the 3D cube. For this step, based on 3D PSF modeling, we first-guess the target location in the IFU by projecting the WCS solution

measured on guider images taken simultaneously. Following Buton et al. (2013), The PSF model is a linear combination of Moffat and Gaussian elliptical profiles. Altogether, the PYSEDM pipeline enables a fully automated spectral extraction of a targeted point source object within five minutes after the end of exposure. This spectrum is typed using SNID and pushed to registered clients, notably the ZTF marshal. If the SNID typing quality value rlap is higher than 4, the SNID type is also pushed.

While no human interaction is required, PYSEDM conveniently allows manual re-extraction or fine-tuning of any step of the pipeline. For instance, one can fine-tune the target location and/or the IFU area considered for 3D PSF extraction in order to avoid bright source contamination.

We show in this paper that SEDM spectra automatically extracted by PYSEDM are relatively flux-calibrated at the few percent level, which enables spectroscopic typing. In the 19 weeks since PYSEDM entered production in early August, we have classified, among other objects, about 400 Type Ia supernovae and 140 Type II supernovae; see classification details in Fremling et al. (in prep.).

To conclude, we would like to emphasize the power of having fully automated spectroscopic facilities in the era of modern time-domain imaging surveys such as ZTF today and LSST in the near future. LSST will operate in the south and will discover up to ten times more transients than ZTF. A fraction of those will be near enough in redshift (say $z < 0.2$) that one could imagine getting a spectrum of most, if not all of them with just a few SEDM-like facitities installed on 2-m class telescopes in the southern hemisphere. PYSEDM has been developed to be extremely flexible and could, therefore, be adapted to any IFU-like spectrograph with very limited effort. Also, while the SEDM very low-resolution ($R \sim 100$) is very efficient for typing, from the perspective of developing new spectroscopic follow-up facilities, we would encourage the use of a higher wavelength resolution design ($R \sim 1000$), high enough to resolve the Hα [NII] structure in order to accurately enable host spectroscopic redshift measurements simultaneously with the transient typing.

References

Aldering, G., Adam, G., Antilogus, P., et al. 2002, Proc. SPIE, 4836, 61
Arribas, S., Mediavilla, E., García-Lorenzo, B., del Burgo, C., & Fuensalida, J. J. 1999, A&AS, 136, 189
Barbary, K. 2016, J. Open Source Softw., 1, 6
Bailey, S., Aldering, G., Antilogus, P., et al. 2009, A&A, 500, L17
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blondin, S., & Tonry, J. L. 2007, ApJ, 666, 1024
Bohlin, R. C., Gordon, K. D., & Tremblay, P.-E. 2014, PASP, 126, 711
Bongard, S., Soulez, F., Thiébaut, É., & Pecontal, É. 2011, MNRAS, 418, 258
Blagorodnova, N., Neill, J. D., Walters, R., et al. 2018, PASP, 130, 035003
Buton, C., Copin, Y., Aldering, G., et al. 2013, A&A, 549, A8
Copin, Y. 2013, HDR Université Claude Bernard - Lyon I
Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, ApJ, 696, 870
Feindt, U., Nordin, J., Rigault, M., et al. 2019, ArXiv e-prints [arXiv:1902.03923]
Filippenko, A. V. 1982, PASP, 94, 715
Fried, D. L. 1966, J. Opt. Soc. Am., 56, 1372
Graham, M. J., Bellm, E. C., Kulkarni, S. R., et al. 2019, PASP, accepted
Ishida, E. E. O., & de Souza, R. S. 2013, MNRAS, 430, 509
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2017, ApJ, 843, 6
Kaiser, N., Aussel, H., Burke, B. E., et al. 2002, Proc. SPIE, 4836, 154
Kissler-Patig, M., Copin, Y., Ferruit, P., Pécontal-Rousset, A., & Roth, M. M. 2004, Astron. Nachr., 325, 159
Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, AJ, 139, 1782
Lantz, B., Aldering, G., Antilogus, P., et al. 2004, Proc. SPIE, 5249, 146
Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. 2016, ApJS, 225, 31
LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
Möller, A., Ruhlmann-Kleider, V., Leloup, C., et al. 2016, JCAP, 12, 008
Piascik, A. S., Steele, I. A., Bates, S. D., et al. 2014, Proc. SPIE, 9147, 91478H
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rigault, M., Brinnel, V., Aldering, G., et al. 2018, A&A, submitted [arXiv:1806.03849]
Sako, M., Bassett, B., Becker, A., et al. 2008, AJ, 135, 348
Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48
Smartt, S. J., Valenti, S., Fraser, M., et al. 2015, A&A, 579, A40
Stone, J. A., & Zimmerman, J. H. 2011, Engineering Metrology Toolbox, 59
Tokovinin, A. 2002, PASP, 114, 1156
Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505