Astronomy & Astrophysics

Wavelet theory applied to the study of spectra of trans-Neptunian objects

A. C. Souza-Feliciano, A. Alvarez-Candal, and Y. Jiménez-Teja

Observatório Nacional, rua General José Cristino 77, 20921-400 Rio de Janeiro, Brazil
e-mail: carolinaastro@on.br

ABSTRACT

Context. Reflection spectroscopy in the near-infrared (NIR) is used to investigate the surface composition of trans-Neptunian objects (TNOs). In general, these spectra are difficult to interpret due to the low apparent brightness of the TNOs, causing low signal-to-noise ratio even in spectra obtained with the largest telescopes available on Earth, making it necessary to use filtering techniques to analyze and interpret them.
Aims. The purpose of this paper is to present a methodology to analyze the spectra of TNOs. Specifically, our aim was to filter these spectra in the best possible way: maximizing noise removal, while minimizing the loss of signal.
Methods. We used wavelets to filter the spectra. Wavelets are a mathematical tool that decompose the signal into its constituent parts, allowing us to analyze the data in different areas of frequencies with the resolution of each component tied to its scale. To check the reliability of our method, we compared the filtered spectra with the spectra of water and methanol ices to identify some common structures between them.
Results. Of the 50 TNOs in our sample, we identify traces of water ices and methanol in the spectra of several of them, some with previous reports, while for other objects there were no previous reports.
Conclusions. We conclude that the wavelet technique is successful in filtering spectra of TNOs.

Key words. kuiper belt: general – techniques: spectroscopic – methods: data analysis

1. Introduction

Trans-Neptunian objects (TNOs) are icy relics left over from the planetary accretion disk that orbit the Sun beyond Neptune. The investigation of their properties, as remnants of the external planetesimal swarms, is essential to understanding their formation and evolution. Investigating their surfaces by spectroscopy is a tool for constraining the chemical properties and the evolution processes that occurred in the solar system (Barucci et al. 2008).

To better understand the objects in this region we rely on observations, mostly ground-based, of their surface properties. In particular, we use reflection spectroscopy in the near-infrared (NIR). In general, the NIR spectra of TNOs are difficult to interpret because of the objects' low apparent brightness, which causes a low signal-to-noise ratio (S/N) even in data obtained with the largest telescopes available on the Earth. Therefore, the spectra of these objects usually have a considerable amount of noise making it necessary to use filtering techniques to analyze and interpret them.

In signal processing, noise is defined as that undesirable, random signal that is mainly introduced by the processes of capture, storage, and transmission of the data. Although inherent to the data and thus difficult to be disentangled from them, the noise must be estimated and removed in the most efficient way without altering the real information contained in the data. Hereafter we use the term signal to indicate that part of the data containing real information, and noise to describe that artificial, random signal introduced by the mechanisms previously described.

Filters are usually applied to remove noise from signals. Due to its random nature, noise is identified with unstructured high frequencies. Filtering techniques try to identify these frequencies and eliminate or attenuate them.

Among the most commonly used techniques are rebinning, the runing box, and the Fourier transform. However, these techniques are not completely efficient because they either decrease the spectral resolution (rebinning and runing box) or they can incorrectly remove high frequency spectral features easily confused with noise (Fourier transform). Two important problems affect this last technique: 1) the definition of "high frequency" is somewhat arbitrary, and lacks a robust definition, and 2) data are processed as a whole. In this sense, the Fourier transform provides the frequencies that globally compose the data. There is no information on the localization of these frequencies, that is, whether they appear throughout the whole signal or just in a certain interval. As they are all treated equally with this global analysis, it is difficult to determine whether a high frequency is truly related to noise or instead a real feature of the signal, thus containing information. This second problem is usually overcome by applying the Fourier transform in windows, but the arbitrary definition of the size of the window and the border effects induced by them made us choose a different technique to filter our TNO data.

Wavelets are similar to the Fourier transform since they decompose the data into the frequencies that compose them, but with the difference that these are localized (although not perfectly). The wavelets perform a multiscale analysis of the data, where the different frequencies that compose the signal are roughly located with an uncertainty that depends on the resolution of the scale. The information provided by the wavelets is thus essentially the same as that from Fourier, but detailed and

split between different scales. With this new approach, a high frequency is not immediately associated with noise, but depending on its modulus (i.e., its presence and weight in the data), it will be considered a significant piece of information or noise.

Wavelets have been used in astrophysics for various purposes: to denoise cosmological simulations (Romeo et al. 2003, 2004), to remove the background light in galaxy clusters and detect faint small objects to complete their luminosity function (Livermore et al. 2017), to identify the different scales in wavelet decomposition with luminous components in groups of galaxies (Da Rocha et al. 2008; Da Rocha & Mendes de Oliveira 2005), and to study mutual events among Galilean satellites (Ramírez et al. 2006), among others.

Here we show how the application of the wavelets to filter the data allows us to reduce the amount of noise in the spectra of 50 TNOs obtained with the instrument SINFONI at the Very Large Telescope in Cerro Paranal, Chile. We describe the sample in the Sect. 2 and the chosen filtering technique in Sect. 3. The methodology used is shown in Sect. 4 and the results are presented in Sect. 5. Finally, the conclusions are drawn in the Sect. 6.

2. Data

All spectra shown in this paper were obtained with the Spectrograph for Integral Field Observations in the Near-Infrared (SINFONI)[1] located in the Cassegrain focus of the unit 4 of the Very Large Telescope (VLT). SINFONI provides a spectroscopic imaging in 3D in the spectral range between 1.05 and 2.45 μm. This instrument has four different modes of observation, J, H, K, and $H+K$, with a spectral resolution of 1500 in the $H+K$ band. The spectra of the sample were obtained using the $H+K$ mode because it allows us to observe the H and K bands simultaneously. The spatial resolution was 0.25 arcsec/spaxels, with corresponding field of view of 8.0 arcsec.

Our sample was selected from the objects observed within an ESO Large Program carried out between 2007 and 2009 (P.I.: M. A. Barucci). The objects analyzed are listed in Table 1. Most of the spectra presented here have already been published. Details of those marked with (+) can be found in Barucci et al. (2011), with (++) in DeMeo et al. (2010), with (−) in Barucci et al. (2010), with (#) in Barucci et al. (2012), with (*) in Guilbert et al. (2009), with (−) in Alvarez-Candal et al. (2007), (**) in Merlin et al. (2007), and Pluto and Charon in Merlin et al. (2010). We note that the spectra of Chiron and 1996 TO_{66} have not been previously published; nevertheless, the details of their reduction can be found in any of the previous references. In this way, according to the classification of Gladman et al. (2008), among the spectra in our sample are included 13 classical objects, 10 resonant objects, 16 centaurs, 4 scattering disk objects, and 5 detached objects.

The sizes of the objects range from 20 to 2700 km (see Table 1), several of which with diameters over 500 km (Pluto, Triton, Haumea, Quaoar, etc.). From a technical point of view, it is easier to detect absorption features in the spectra of the largest objects, due to their S/N and also because large objects can retain a larger reservoir of volatiles (Schaller & Brown 2007). Smaller objects, which constitute most of our sample, are more difficult to deal with because (1) they are too small to have kept the volatiles that were accreted onto them when they were formed (Schaller & Brown 2007); and/or (2) of their small sizes, their

apparent brightness is too faint and therefore the S/N is too low to detect small absorption features.

Unfortunately, nothing can be done in case (1), but we can tackle case (2) and try to work with the spectra generated by these objects through the application of techniques that improve their signal-to-noise ratios until the arrival of the next generation of telescopes. Therefore, we applied the wavelets to filter all objects in our sample. For large TNOs ($d > 500$ km) (Brown 2012), especially those that have been well studied in the literature, the goal was to verify whether the application of wavelets can be a tool for reducing the noise and for recognizing the absorption bands already reported in the literature without losing information. As the result was positive, we applied the same technique to smaller TNOs ($d < 500$ km) and tried to identify possible evidence of absorption bands of water and methanol ices that the noise could be hiding.

3. Data processing

As mentioned above, due to the overall faintness of the TNOs, their spectra usually have relatively low S/N (seldom over 20 or 30). It is therefore necessary to use a filtering technique to analyze and interpret them.

The main purpose of our study is to use the wavelets technique to reduce the maximum amount of noise without changing the shapes and positions of the possible absorption bands of the spectra, if present, in order to identify them. Once filtered, some spectral parameters are calculated to estimate the possibility of absorption detection, due to some ices of astrophysical interest such as water and methanol ices. After that, these spectra are compared to spectra calculated from optical constants measured in the laboratory for the ices mentioned. This step is necessary to link the absorptions detected in the filtered spectra with the ices we are looking for.

3.1. Wavelets

The wavelet transform is a multiresolution technique which decomposes the data into their constituent frequencies, roughly located in time (or wavelength) and frequency simultaneously. It is based on the so-called wavelet function or mother wavelet ψ, which can be any squared-integrable function with zero mean:

$$\int_{-\infty}^{+\infty} \psi(t)\mathrm{d}t = 0. \tag{1}$$

These two simple properties guarantee that ψ has an undulating shape, and is the reason why it is called a wavelet. For each wavelet ψ, we can build its associated wavelet family, consisting of the original wavelet function ψ dilated and translated by different factors, the scale j and the translation factor k, respectively:

$$\psi_{j,k} = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t - 2^j k}{2^j}\right). \tag{2}$$

It can be proved that these wavelet families are mathematically orthonormal bases of the $\{L\}^2$-space of the squared integral functions. This means that any function (or data), provided that it is smooth enough, can be decomposed as the linear combination of the elements of the wavelet family

$$f(t) = \sum_{j=0, j \in \mathbb{Z}}^{\infty} \sum_{k \in \mathbb{Z}} d_{j,k}\psi_{j,k}(t), \tag{3}$$

Table 1. Spectral parameters of observed objects.

Number	Wavelet	Scal	$D_{2.0\,\mu m}$ (%)	$D_{2.27\,\mu m}$ (%)	Diameter (km)[1]	S/N_o	S/N_w
2060 Chiron	Bior 5.5	7	3.7 ± 0.07	13.3 ± 1.3	215.6	56.1	357.6
5145 Pholus +	Rbio 3.9	5	7.7 ± 0.4	33.9 ± 9.2	119	8.5	38.2
10199 Chariklo *	Bior 5.5	5	6.8 ± 0.2	1.84 ± 0.09	231	86.2	179.5
15874 1996 TL_{66} *,+	Bior 5.5	8	22.7 ± 12.8	35.1 ± 16.6	339	6.3	46.6
19308 1996 TO_{66}	Bior 5.5	7	33.6 ± 40.1	33.0 ± 37.0	200	0.95	9.91
26375 1999 DE_9 *	Bior 5.5	8	9.22 ± 0.05	5.36 ± 1.04	311	10.6	213.0
28978 Ixion *	Rbio 3.9	7	11.5 ± 1.5	27.9 ± 16.8	617	16.5	92.2
32532 Thereus *	Bior 5.5	8	10.0 ± 1.0	8.9 ± 1.9	62	15.7	122.2
38628 Huya –	Bior 5.5	6	7.38 ± 0.43	21.0 ± 2.4	458	9.0	31.6
42355 Typhon *	Bior 5.5	7	14.5 ± 0.2	4.8 ± 0.8	185	18.5	81.2
44594 1999 OX_3 +	Rbio 3.5	8	4.15 ± 0.85	1.46 ± 0.43	135	15.7	270.0
47171 1999 TC_{36} *	Rbio 2.6	7	8.16 ± 0.74	0.74 ± 0.12	393	24.7	185.0
47932 2000 GN_{171} *	Rbio 2.8	6	20.0 ± 4.0	21.7 ± 7.40	147	9.08	34.68
50000 Quaoar *	Rbio 2.8	8	28.1 ± 0.4	11.8 ± 0.5	1036	23.7	28.9
52872 Okyrhoe ++	Sym 15	8	2.57 ± 0.07	0.72 ± 0.01	35	56.7	180.6
54598 Bienor *	Rbio 2.4	8	12.7 ± 1.9	25.4 ± 12.8	198	6.4	96.5
55565 2002 AW_{197} *	Db 2	8	13.8 ± 7.9	21.3 ± 6.4	768	16.4	190.2
55576 Amycus +	Bior 5.5	7	12.6 ± 0.4	13.5 ± 2.0	104	15.2	148.8
55637 2002 UX_{25} +	Bior 5.5	7	12.4 ± 1.3	38.0 ± 8.8	692	7.0	11.6
55638 2002 VE_{95} #	Bior 5.5	8	10.28 ± 0.05	11.3 ± 1.1	250	21.4	267.4
60558 Echeclus *	Bior 5.5	8	2.5 ± 0.4	1.2 ± 0.5	65	10.1	41.8
73480 2002 PN_{34} ++	Bior 5.5	8	13.9 ± 2.2	2.0 ± 0.6	112	12.0	95.7
83982 Crantor *	Db 18	8	14.0 ± 1.8	34.3 ± 11.0	59	15.8	141.9
90377 Sedna –	Bior 5.5	7	25.7 ± 10.0	30.7 ± 6.2	906	12.5	108.6
90482 Orcus *	Rbio 2.6	8	38.3 ± 1.3	19.6 ± 1.6	958	15.9	22.5
90568 2004 GV_9 *	Rbio 2.8	7	9.1 ± 0.2	26.2 ± 0.5	680	14.2	101.3
95626 2002 GZ_{32} +	Rbio 2.4	7	8.27 ± 1.25	4.87 ± 1.38	237	11.5	433.8
119951 2002 KX_{14} *	Rbio 3.7	8	7.9 ± 0.1	34.9 ± 0.7	455	7.0	290.5
120061 2003 CO_1 +	Rbio 2.8	7	5.16 ± 0.33	0.42 ± 0.21	94	29.9	52.3
120132 2003 FY_{128} *	Sym 13	8	28.0 ± 6.0	10.8 ± 13.7	460	3.0	18.0
120348 2004 TY_{364} +	Bior 5.5	8	2.27 ± 0.02	0.60 ± 0.01	512	16.0	146.8
134340 Pluto	Bior 3.1	6	4.31 ± 0.12	76.0 ± 17.6	2372	46.1	46.2
136108 Haumea **	Rbio 3.9	5	59.8 ± 1.8	27.0 ± 2.6	1240	13.0	14.7
136199 Eris *	Bior 5.5	6	11.6 ± 0.8	92.4 ± 11.6	2326	18.46	21.2
144897 2004 UX_{10} +	Rbio 2.8	6	18.5 ± 3.3	34.6 ± 6.9	398	11.4	63.6
145451 2005 RM_{43} +	Bior 5.5	8	21.9 ± 1.8	22.0 ± 1.9	252	5.4	36.1
145452 2005 RN_{43} *	Bior 5.5	8	1.08 ± 0.06	0.94 ± 0.01	679	18.8	126.5
145453 2005 RR_{43} +	Bior 5.5	8	76.4 ± 2.6	17.6 ± 0.3	670	3.6	5.2
174567 Varda +	Db 20	8	7.0 ± 0.5	9.9 ± 1.5	792	12.5	128.5
208996 2003 AZ_{84} +,*	Db 4	8	20.2 ± 3.3	16.1 ± 6.7	727	7.7	41.0
229762 2007 UK_{126} +	Rbio 2.6	8	11.9 ± 1.1	10.8 ± 3.1	599	10.5	95.1
250112 2002 KY_{14} +	Bior 5.5	8	5.0 ± 0.3	7.9 ± 0.7	47	19.8	110.1
281371 2008 FC_{76} +	Bior 5.5	8	2.88 ± 0.02	13.3 ± 2.8	68	10.9	231.7
307616 2003 QW_{90} *	Bior 5.5	8	3.85 ± 0.05	12.2 ± 1.3	366	2.05	11.0
309737 2008 SJ_{236} +	Bior 5.5	7	12.5 ± 0.3	63.0 ± 52.5	18	5.6	71.8
455502 2003 UZ_{413} +	Bior 5.5	8	25.0 ± 2.2	2.6 ± 1.8	1214	7.0	130.7
2007 UM_{126} +	Rbio 3.7	7	7.4 ± 0.2	19.4 ± 6.4	42	10.4	130.0
2007 VH_{305} +	Rbio 2.8	6	13.2 ± 4.6	23.9 ± 13.8	24	5.6	39.4
Charon	Rbio 2.8	8	53.4 ± 0.7	3.2 ± 0.1	1212	9.8	9.8
Triton[2]	Bior 4.4	6	16.6 ± 1.1	42.4 ± 3.0	2706	51.4	54.0

Notes. This table give us an idea about the presence of methanol and/or water ices on the surface of theses objects. We took advantage of the last two columns to show the signal-to-noise ratio before and after the filtering, and in Col. 2 we show which wavelet was used to perform the filtering, where Bior is the wavelet Biorthogonal, Rbio is Reverse biorthogonal, Sym is Symlets, and Db is Daubechies. [1] The diameters shown here were taken from the survey "TNOs are cool" (http://public-tnosarecool.lesia.obspm.fr/). [2] Its believed that Triton had its origin in the trans-Neptunian region, and that it was captured by Neptune during the evolution of the solar system.

where the coefficients $d_{j,k}$ are the wavelet coefficients. These coefficients are calculated according to the following expression:

$$d_{j,k} = \int_{-\infty}^{+\infty} f(t)\overline{\psi_{j,k}(t)}\mathrm{d}t. \qquad (4)$$

The wavelet coefficients are related to the details (frequencies) that compose the data. Their lower scales (j) give information on the finer details that compose the data, such as high frequency structures, pseudo-impulsive features, or noise. These fine details are very well localized in time due to the small support size of the associated wavelet functions $\psi_{j,k}$ (its width in the time space). Higher j scales tell us about coarser details, low frequency structure in the data, usually related with information. Due to the increase in the support size of the $\psi_{j,k}$, these coarse details are localized in time with a higher uncertainty. However, the support size of a function in time is inversely proportional to its support size in the Fourier domain. That means that the narrower the $\psi_{j,k}$ are, the wider their support in frequency becomes, thus increasing the uncertainty in frequency. For this reason, in the wavelet decomposition the high frequencies that compose the data are not very well identified, although exceptionally well localized in the data (i.e., we know where they are, but not who they are), whereas the low frequency structures are better defined, but we do not know exactly where they appear (i.e., we know who they are, but not where they are). So, if Fourier transform provides the exact decomposition in frequency of the data independently of its localization, the wavelet transform tells us approximately which frequencies compose the data and approximately the intervals where they appear. We lose precision in frequency to gain some information on time.

This is very useful for disentangling the noise from the signal. In the wavelet theory, the noise is no longer related to the high frequencies, but to those structures that do not have a strong weight in the decomposition in a certain time interval. For example, the high frequency components of a narrow absorption line would be kept as signal (information) since these high frequencies contain structural information of the data in that interval. However, these same high frequencies can be identified as noise in any other interval of the data where the subjacent structure is smoother, like a broad line in a noisy spectrum, since the weight of these high frequencies is small compared to that of the low frequencies. Thus, the noise will be defined as the high frequency in a certain interval which does not have a relative importance for the structure of the data in that interval, identified by the modulus of its associated wavelet coefficient $d_{j,k}$. To quantify this idea, Donoho & Johnstone (1994) proved that when there is Gaussian noise, there is a certain threshold (called the universal threshold) below which a wavelet coefficient has a very high statistical probability of carrying noise. This universal threshold for a signal of size N polluted with white noise of variance σ^2 is defined as

$$T = \sigma \sqrt{2 \log N}. \qquad (5)$$

The only problem of the universal threshold is that it assumes the variance of the noise to be known. As this is not usually the case with real data, it can be proved that the following parameter is a robust estimator of the variance σ^2 of the noise,

$$\sigma = \frac{M}{0.6745}, \qquad (6)$$

where M is the median of the absolute moduli of the wavelet coefficients in the first scale.

Thus, to filter our noisy TNO spectra we decompose them up to a certain scale j using a certain wavelet family, calculate the corresponding universal threshold T, and set to zero all the coefficients with $|d_{j,k}| < T$. This is called a "hard filter" and it guarantees that the characteristics of the signal, in our case the shape and depth of the absorption bands in our spectra, are preserved after removing the noise.

3.2. Selection of the wavelet

There are several wavelet functions (each with its associated wavelet family) that lead to different wavelet decompositions. The wavelets functions can be real or imaginary, have peaks or be smooth, be symmetric or asymmetric, be more or less oscillating, etc. Therefore, there is no "perfect" wavelet function; we must choose the one that works best for our problem according to two main criteria, the characteristics of the data and our final purpose.

As the wavelet decomposition is just the result of the convolution of the data with the mother wavelet, dilated and translated by different factors, wavelet functions that are similar to the data perform better. In this sense, if our data have symmetric features (e.g., the absorption lines in our spectra) a symmetric wavelet is preferred. A smooth, not very oscillating wavelet would also be desirable in our case, since the spectra do not present sharp peaks or highly oscillating features (e.g., a seismic signal).

The final goal of the wavelet decomposition is also very important for the selection of the wavelet. Wavelets can be used to detect singularities in the data, for instance sharp peaks, extrema, inflection points (see, e.g., Ramírez et al. 2006). For this purpose, the support size of the wavelet is crucial since wavelets with wide support introduce too much uncertainty in time, and the singularity will not be tightly identified. When we are interested in filtering the data, the support size of the wavelet is not essential (although a narrow wavelet function is always desirable), but its compression capacity is much more important. The wavelet transform keeps the energy of the data, meaning that the norm of the signal is the same as the norm of the wavelet coefficients of its decomposition. As the filtering algorithm is based on the comparison of the modulus of the wavelet coefficients carrying noise with the modulus of the coefficients associated with the signal, a large difference between these two sets would be convenient. If the energy of the signal were split into many coefficients, this comparison might not be clear and a coefficient containing signal could be misidentified as noise. However, if the energy of the information were concentrated in just a few wavelet coefficients, the filtering process would be straightforward. The concept of vanishing moments helps us know the capacity of concentration (compactness of the decomposition) that has the wavelet. In practice, the longer the number of vanishing moments that the function has, the fewer the coefficients needed in the decomposition (Eq. (3)).

Therefore, to select the wavelet that best filters our data, up to a certain scale j, we can follow two paths: (i) perform the filtering for different wavelets and different scales for each input spectrum one by one and decided manually which is the best or (ii) devise an automatic algorithm. The first case might work for one or two spectra, but it is highly inefficient for a larger number, therefore we will follow path (ii).

We devised a blind test in order to scan among the discrete wavelets offered by the PyWavelet package of Python with changing scales between 1 and 8. As a test sample, we used 41 spectra of water ice with signal-to-noise ratios ranging from about 1 to about 200, with emphasis in the range 1–50.
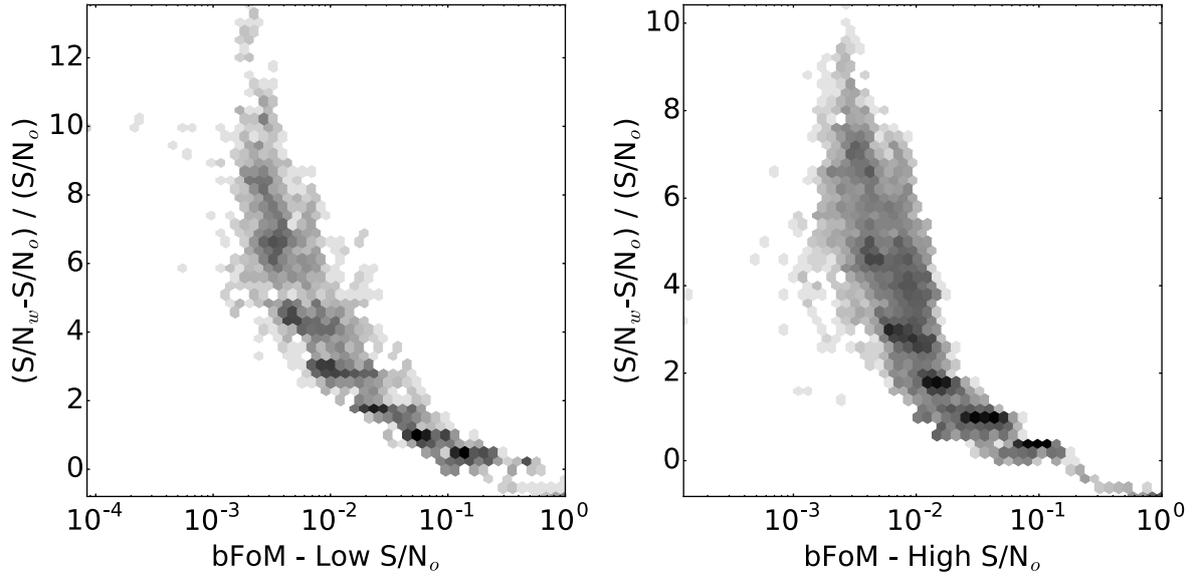
Fig. 1. Maps of the distribution of the bFoM compared to the relative gain in S/N_w. The *left panel* shows data with $S/N_o < 20$, while the *right panel* shows data with $S/N_o > 20$. (See text.)

In order to select the optimal wavelet we use a quadratic sum of three different figures of merit (FoMs, taken from Meza et al. 2010). All of them are blind, i.e., with no a priori information regarding the data (e.g., what ice lies underneath the noise.)

Before continuing we define our input spectrum (data) as $I = [I_i, ..., I_M]$ and the filtered spectrum (signal) as $F = [F_i, ..., F_M]$.

The first FoM is the roughness (adapted from Hayat et al. 1999),

$$\rho = \sum_{i=1}^{M} |[-1, 1] * F_i| / \sum_{i=1}^{M} |F_i|, \tag{7}$$

which measures the roughness of a signal. The asterisk stands for convolution. A constant, smooth signal will have $\rho = 0$, while increasing pixel-to-pixel variation will have larger values of ρ.

A second FoM is the noise reduction (NR, adapted from Chen et al. 2003),

$$NR = N_0/N_1, \tag{8}$$

where N_0 is the average of the power spectrum of the data and N_1 is the average of the power spectrum of the signal. The best filtering is obtained when NR is high.

The last FoM we use is the residual non-uniformity (RNU, adapted from Wang et al. 2008),

$$RNU = 100/\overline{F} \sqrt{(1/M \times \sum_{i=1}^{M} (F_i - \overline{F})^2)}, \tag{9}$$

where we search for the most uniform filtered signal, and therefore for low values of RNU. The overlined quantities are mean values.

In order to define our metric we first invert NR to search for its minimum and renormalize all FoMs to the interval [0,1]. Then we define our blind figure of merit (bFoM) as

$$bFoM = \sqrt{\sum_{i=1}^{3} (FoM_i)^2/3}; \tag{10}$$

as such, the optimal wavelet and scale for filtering will be given by the minimum of bFoM.

By construction we know the input signal-to-noise ratio, S/N_o, of the spectra in the test sample. Each one was filtered with 58 different wavelets, in eight different scales, amounting to 454 filtered spectra per input spectrum. Considering that we have 41 modeled spectra, in the end we have 19014 filtered spectra each with a value of bFoM. To help visualize the results we created maps that show a clear trend to larger gains in S/N with decreasing values of bFoM (Fig. 1). Therefore, we adopt as our criterion for selecting the wavelet and scale the one that minimizes bFoM; as such, we are selecting the wavelet and scale that optimally balances the smoothness, uniformity, and noise removal of the data. Upon visual inspection of the signal (the filtered spectrum), the results are satisfactory.

It must be mentioned that although low values of bFoM ensure good gains in S/N, it does not necessarily maximizes it; this is a trade off between knowing a priori what the "noiseless" spectrum should look like and no a priori assumptions.

Figure 1 has been split into two panels. The left panel shows the maps for model spectra with $S/N_o < 20$, while the right panel does it for $S/N_o > 20$. We have chosen this value as the average S/N of the spectra in our sample (see Table 1) and to illustrate that the performance of bFoM is not correlated by the S/N_o of the spectrum.

We also computed the performance of a typical Fourier filtering, applying a fast Fourier transform (S/N_f), against the filtering with the optimal wavelet (S/N_w). It can be seen in Fig. 2 that the optimal wavelet clearly outperforms the Fourier filtering even for low S/N_o spectra. It is also clear that S/N_w is at least a factor of 2 better than the S/N_o.

Therefore, in what follows, we always use the wavelet and scale that minimizes the bFoM and perform a visual inspection to corroborate the quality of the filtering.

4. Methodology

We applied the algorithm mentioned above to our 50 spectra. The signal-to-noise ratio before and after the filtering are shown in Table 1.
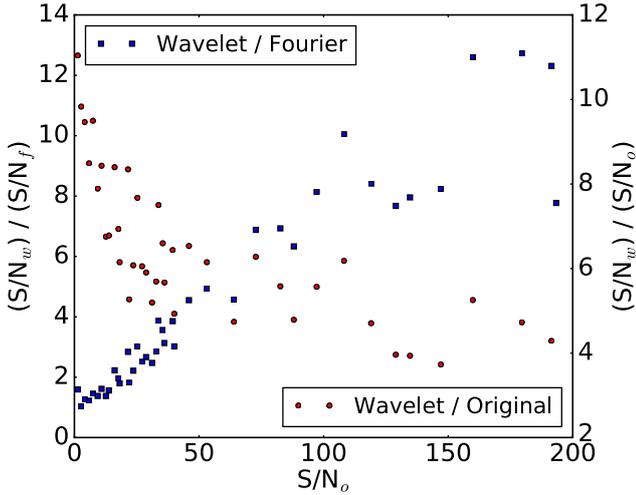
Fig. 2. Comparison between the performances of the optimal wavelet, found by our algorithm, and a Fourier filtering. The blue squares show the ratio of S/N between wavelet and Fourier filtering, while the red circles show the S/N between the wavelet filtering and the original data.

In the filtered spectra we measured two absorption band depths, the first at 2.0 μm, which gives an idea about the presence of water ice, and the second at 2.27 μm related to methanol (e.g., Merlin et al. 2012). Fluxes were computed as the median value between 1.98–2.02, 2.18–2.22, and 2.25–2.29 μm. The error assigned to each value comes directly from the standard deviation of the points used to compute the depth of the absorption band in each interval. These results are presented in Table 1.

Finally, we compare the filtered spectra with laboratory ices of water and methanol ices, obtained using the Hapke (1981) model and optical constants from Quirico & Schmitt (1997). The objective was to confirm the presence of these ices in the objects where their presence was reported, in order to validate our technique. We also looked for evidence of these ices in the spectra of TNOs that were previously neglected because of their low signal-to-noise ratios.

5. Results

5.1. Filtering

We show here two examples that are representative of our sample: a large and a small body (according to the classification of Brown 2012).

Figure 3 shows the filtering applied to the spectrum of 136108 Haumea, which is a large TNO with an apparent brightness high enough to generate a spectrum with a high signal-to-noise ratio. According to the algorithm, the wavelet that best filters this spectrum is the wavelet Reverse Bi-orthogonal 3.9 in the scale 5. The technique was successful in filtering the spectrum because the positions of the absorption bands of the original spectrum were not displaced nor were their shapes altered. Analyzing the residue distribution, we see that there are no structures or patterns, and the signal-to-noise ratio increased. Zooming into the region between 1.9 and 2.3 μm (Fig. 4) where there is absorption due to water ice, it is possible to perceive with more clarity that the technique did not alter the shape of the band. This indicates that the wavelets only removed noise while respecting the signal of the spectrum.
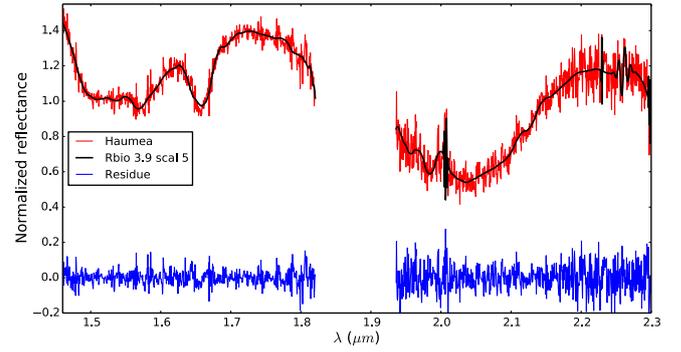

Fig. 3. Spectrum of 136108 Haumea. The original data is shown in red, the filtered spectrum with the wavelet RBio 3.9 (scale 5) is shown in black, and the difference between the original and filtered spectra is shown in blue.
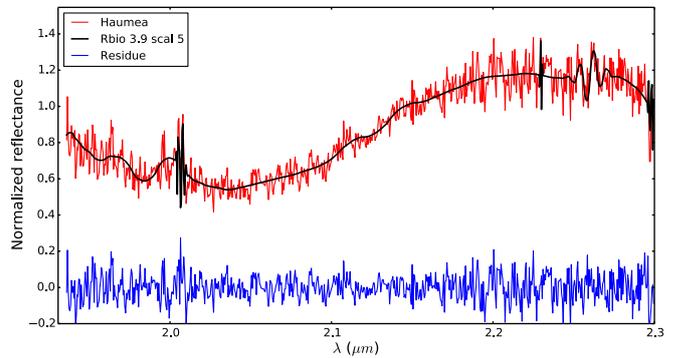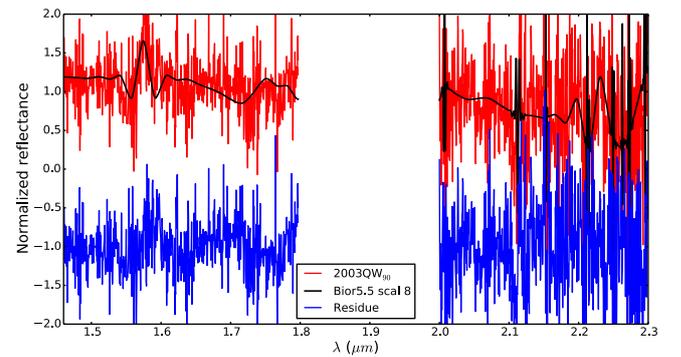

Fig. 4. A zoom-in of the region between 1.9 and 2.3 μm of the Haumea spectrum for better visualization of the filtration. The original spectrum (before filtering) is shown in red, the filtered spectrum with the wavelet RBio 3.9 (scale 5) is shown in black, and the difference between the original and filtered spectra is shown in blue.

Fig. 5. Spectrum of 307616 2003 QW$_{90}$. The original data is shown in red, the filtered spectrum with the wavelet Bior 5.5 (scale 8) is shown in black, and the difference between the original and filtered spectra is shown in blue. The residue distribution was shifted downward by 1 unit for better visualization.

Figure 5 shows how the technique works on the spectra of smaller TNOs, for instance the spectrum of 2003 QW$_{90}$. In this case the optimal scale for filtering according to the algorithm is the Bi-orthogonal 5.5 in scale 8. Analyzing Fig. 5, we see that although the amount of noise is large, the technique did not insert any obvious artifact when trying to filter the spectrum. The amount of noise decreased, but some noise still remains in the spectrum. We conclude that the wavelet simple thresholding
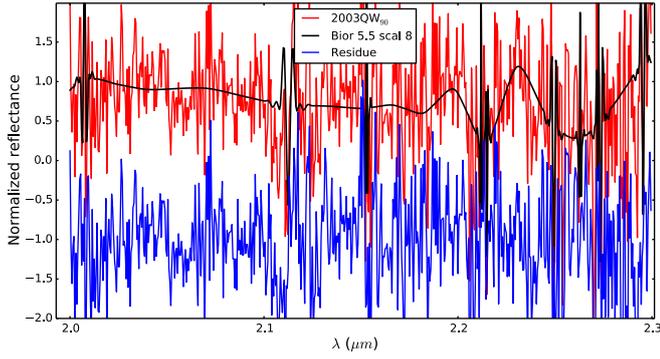
Fig. 6. Zoom-in of the region between 1.9 and 2.3 μm of TNO 2003 QW$_{90}$. The original data is shown in red, the filtered spectrum with the wavelet Bior 5.5 (scale 8) is shown in black, and the difference between the original and filtered spectra is shown in blue. The residue distribution was shifted downward by 1 unit for better visualization.

has a limit: it filters to the point where it is no longer possible to remove noise without removing signal from the spectrum. Analyzing the bottom box, we see that the residue is larger than in the previous case, but there are still no obvious structures or patterns in it, as can be seen in Fig. 6 where we show a zoom of the final region of the spectrum of this TNO.

5.2. Limitations

It can be seen that, whenever the spectra have high S/N$_o$, the absorption structures are very well respected. In the case of lower S/N$_o$ this might not strictly be the case because the wavelets' optimal performance assume a signal that is contaminated by white (Gaussian) noise.

Our data was obtained by photon counting, and is thus contaminated by Poisson noise. This is not a big problem since in the limit of large numbers the Poisson distribution tends to the Gaussian (see Bevington & Robinson 2003). More complicated are systematic noises (instrumental and calibration), which do not necessarily follow a statistical distribution and will affect any filtering technique. Examples of this can be seen in Figs. 1 and 2 with a feature close to 2.2 μm due to incomplete removal of telluric features.

In summary, through this analysis we conclude that the application of the wavelets technique worked well to filter out TNO spectra in both cases (low and high signal-to-noise ratios); therefore, we will assume it works just as well for the intermediate cases. Now, we show the results of the comparisons between these de-noised spectra and the spectra of water and methanol ices in order to find evidence of these ices on the surfaces of these objects.

5.3. Comparison

5.3.1. Water ice

According to Table 1 the candidates in our sample that possess water ice are Quaoar, Orcus, 2003 FY$_{128}$, Haumea, Sedna, 2005 RR$_{43}$, and Charon (all with D$_{2.0}$ > 25%). With shallower bands we have 2003 AZ$_{84}$, 2005 RM$_{43}$, 1996 TL$_{66}$, 2004 UX$_{10}$, 2003 UZ$_{413}$, Triton, Crantor, and Huya (15 < D$_{2.0}$ < 25%). Among these, those that showed possible evidence through direct comparison with the water ice spectrum were Quaoar, Orcus, Haumea, 2005 RR$_{43}$, and Charon. The amount of noise remaining in the spectra of the other TNOs did not allow the identification of the possible water ice band. However, the drop
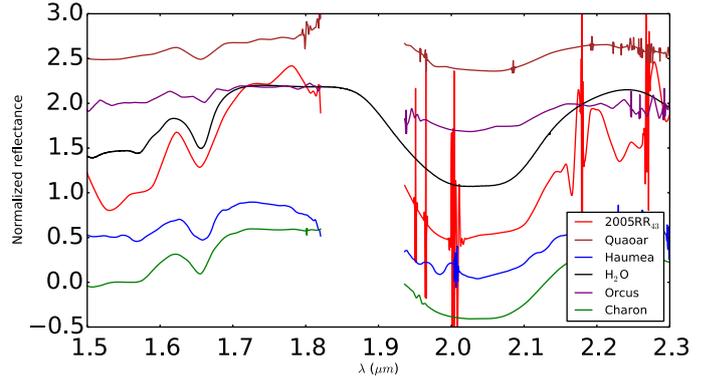

Fig. 7. Normalized reflectance vs. wavelength. The spectra were shifted in flux for better visualization. The region close to 1.85 μm was masked due to the absorptions of the atmosphere. In green the spectrum of Charon, in blue Haumea, in red 2005 RR$_{43}$, in purple Orcus, and in brown, Quaoar. There is a coincidence of the band at 1.65 and 2 μm in all spectra of TNOs, which are signatures of the water ice in these objects.
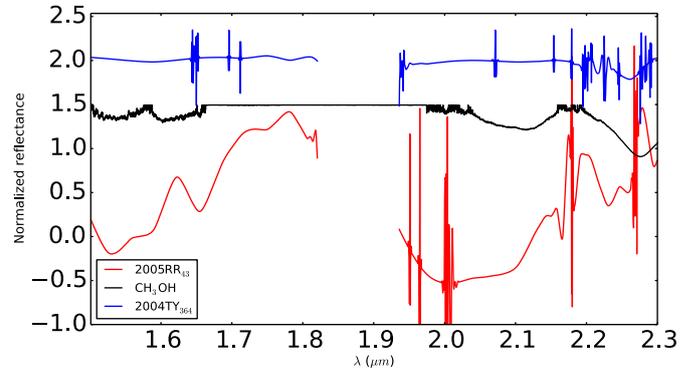

Fig. 8. Normalized flux vs. wavelength. The spectra were shifted in flux for better visualization. From top to bottom the spectra shown are 2004 TY$_{364}$ in blue, the model of methanol (CH$_3$OH) in black, and 2005 RR$_{43}$ in red.
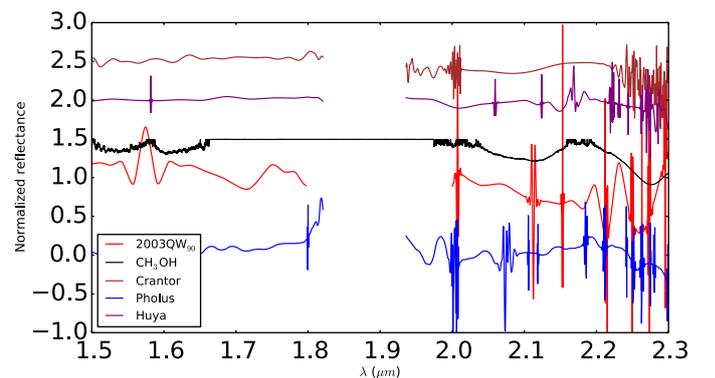

Fig. 9. Normalized flux vs. wavelength. The spectra were shifted in flux for better visualization. From top to bottom, the spectra shown are Crantor in brown, Huya in purple, the model of methanol (CH$_3$OH) in black, 2003 QW$_{90}$ in red, and Pholus in blue.

in flux in the 2 μm region is indicative that some absorption is occurring in the region. Figure 7 shows spectra of objects displaying clear water ice features.

For large TNOs we note the similarity between the water ice spectrum and the object spectra. The band at 1.65 μm is conspicuous in all spectra. This band indicates the presence of

crystalline structure ice on the surface of TNOs, which might be related to a recent exposition of the ice that has had no time to become amorphous due to space weathering. The band at 2 μm appears in all spectra, but with different depths.

5.3.2. Methanol ice

The objects in our sample with evidence of having methanol ice, according to the criterion proposed in Merlin et al. (2012) and after comparison with the spectrum obtained in laboratory to methanol ice, are 2004 TY$_{364}$, 2003 QW$_{90}$, Pholus, Huya, Crantor, and 2005 RR$_{43}$. For the other objects in our sample it was not possible to find structures that could indicate the presence of this ice. In Fig. 8 we show the larger objects in the sample, while in Fig. 9 we show the smaller objects with possible presence of methanol in their surface composition.

Our analysis focuses on the final part of the spectrum, specifically the region beyond 2 μm. In this region the amount of noise is large and therefore is it difficult to detect methanol ice structures for any size range. However, it is possible to perceive signs as flux drop that could be associated with small and shallow absorption features. Nevertheless, care must be taken not to misinterpret some spurious features, as mentioned above.

6. Conclusions

In this paper, we show how the application of the wavelets allowed us to analyze the spectra of 50 trans-Neptunian objects with diameters between 50 and 2300 km. Some of the objects in our sample already have data available in the literature and were taken as validators of the technique. As the results were positive, we applied the technique for the rest of the objects in our sample.

We performed the filtering of the spectra of the sample in a two-step process. First we developed an algorithm that determines the best wavelet and the respective scale that optimizes the result of the filtering, and we filtered the spectra with the wavelet and scale indicated by the algorithm. After filtration, we calculated some spectral parameters to find out in which of the sample spectra there was the possibility of finding structures similar to those found in water and methanol ice spectra. We compared the filtered spectra of the candidates with the spectra calculated from optical constants measured in the laboratory for water and methanol ices. According to the similarity between the spectra, we were able to recognize the absorption bands associated with the water ice and to identify some evidence of the presence of methanol ice in some of the sample spectra. These results are summarized in Table 2.

We consider that the application of wavelets to filter the spectra was successful because it was possible to analyze TNO spectra that had previously been considered unusable due to the amount of noise they had.

We checked whether there was any relation between dynamical grouping (populations) and the possible existence of ices. We did not find any result of relevance, due to the small number statistics. Future developments will have to wait until new data becomes available, especially for smaller objects, when we will be able to apply filtering by wavelets to extract the maximum information possible.

Table 2. Summary of our results: the TNOs with the respective ices.

Name	Ices
5145 Pholus	Methanol+
38628 Huya	Methanol*
50000 Quaoar	Water+
83982 Crantor	Methanol+
90482 Orcus	Water+
120348 2004 TY$_{364}$	Methanol+
136108 Haumea	Water+
145453 2005 RR$_{43}$	Water+, Methanol*
307616 2003 QW$_{90}$	Methanol*
Charon	Water+

Notes. The ices marked with + have detections in the literature, and those with * need confirmation.

References

Alvarez-Candal, A., Barucci, M., Merlin, F., Guilbert, A., & de BerghC.2007, A&A, 475, 369
Barucci, M. A., Brown, M. E., Emery, J. P., & Merlin, F. 2008, The Solar System Beyond Neptune (Tucson: University of Arizona Press), 143
Barucci, M., Dalle Ore, C. M., & Alvarez-Candal, A., et al. 2010, AJ, 140, 2095
Barucci, M. A., Alvarez-Candal, A., & Merlin, F., et al. 2011, Icarus, 214, 297
Barucci, M. A., Merlin, F., & Perna, D., et al. 2012, A&A, 539, A152
Bevington, P., & Robinson, D. 2003, Data Reduction and Error Analysis for the Physical Sciences, (New York: McGraw-Hill Higher Education)
Brown, M. E. 2012, Annu. Rev. Earth Planet. Sci., 40, 467
Chen, J., Shao, Y., Guo, H., Wang, W., & Zhu, B. 2003, IEEE Trans. Geosci. Remote Sens., 41, 2119
Da Rocha, C., & Mendes de Oliveira, C. 2005, MNRAS, 364, 1069
Da Rocha, C., Ziegler, B. L., & Mendes de Oliveira, C. 2008, MNRAS, 388, 1433
DeMeo, F., Barucci, M., & Merlin, F., et al. 2010, A&A, 521, A35
Donoho, D. L., & Johnstone, I. M. 1994, Biometrika, 81, 425
Gladman, B., Marsden, B. G., & VanLaerhoven, C. 2008, in The Solar System Beyond Neptune (Tucson: University of Arizona Press), 43
Guilbert, A., Alvarez-Candal, A., & Merlin, F., et al. 2009, Icarus, 201, 272
Hapke, B. 1981, J. Geophys. Res.: Solid Earth, 86, 3039
Hayat, M. M., Torres, S. N., Armstrong, E., Cain, S. C., & Yasuda, B. 1999, Appl. Opt., 38, 772
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, APJ, 835, 113
Merlin, F., Guilbert, A., & Dumas, C., et al. 2007, A&A, 466, 1185
Merlin, F., Barucci, M., & de Bergh C., et al. 2010, Icarus, 210, 930
Merlin, F., Quirico, E., Barucci, M., & De Bergh C. 2012, A&A, 544, A20
Meza, P., San Martin, C., Vera, E., & Torres, S. 2010, in Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (Berlin, Heidelberg: Springer), 285
Quirico, E., & Schmitt, B. 1997, Icarus, 127, 354
Ramírez, M., Jiménez, Y., González, M., Berrocoso, M., & Sánchez-Francisco, M. 2006, A&A, 448, 1197
Romeo, A. B., Horellou, C., & Bergh, J. 2003, MNRAS, 342, 337
Romeo, A. B., Horellou, C., & Bergh, J. 2004, MNRAS, 354, 1208
Schaller, E. L., & Brown, M. 2007, ApJ, 659, L61
Wang, B., Liu, S., & Bai, L. 2008, Opt. Commun., 281, 2040