The Astrophysical Journal Letters, 881:L9 (7pp), 2019 August 10 https://doi.org/10.3847/2041-8213/ab3581

How to Find Variable Active Galactic Nuclei with Machine Learning

Andreas L. Faisst[1], Abhishek Prakash[1], Peter L. Capak[1,2], and Bomee Lee[1]
[1] IPAC, California Institute of Technology 1200 E California Boulevard, Pasadena, CA 91125, USA; afaisst@ipac.caltech.edu, aprakash@ipac.caltech.edu
[2] Cosmic Dawn Center (DAWN), Copenhagen, Denmark
Received 2019 June 24; revised 2019 July 23; accepted 2019 July 24; published 2019 August 6

Abstract

Machine-learning (ML) algorithms will play a crucial role in studying the large data sets delivered by new facilities over the next decade and beyond. Here, we investigate the capabilities and limits of such methods in finding galaxies with brightness-variable active galactic nuclei (AGNs). Specifically, we focus on an unsupervised method based on self-organizing maps (SOM) that we apply to a set of nonparametric variability estimators. This technique allows us to maintain domain knowledge and systematics control while using all the advantages of ML. Using simulated light curves that match the noise properties of observations, we verify the potential of this algorithm in identifying variable light curves. We then apply our method to a sample of ~8300 *WISE* color-selected AGN candidates in Stripe 82, in which we have identified variable light curves by visual inspection. We find that with ML we can identify these variable classified AGN with a purity of 86% and a completeness of 66%, a performance that is comparable to that of more commonly used supervised deep-learning neural networks. The advantage of the SOM framework is that it enables not only a robust identification of variable light curves in a given data set, but it is also a tool to investigate correlations between physical parameters in multidimensional space—such as the link between AGN variability and the properties of their host galaxies. Finally, we note that our method can be applied to any time-sampled light curve (e.g., supernovae, exoplanets, pulsars, and other transient events).

Key words: galaxies: active – galaxies: evolution – galaxies: photometry – methods: data analysis

1. Introduction

Over the next decade, new facilities will deliver a tremendous amount of data to study astrophysical phenomena. In order to trawl through these large data volumes, fast, automated, and efficient methods are needed. Machine-learning (ML) algorithms are a powerful tool to identify, classify, characterize, and visualize astronomical objects and the correlations of their physical properties in multidimensional parameter space. They are already being used to derive photometric redshifts and other physical properties of galaxies (Masters et al. 2015; Krakowski et al. 2016; Speagle & Eisenstein 2017a, 2017b; Siudek et al. 2018; Bonjean et al. 2019; Davidzon et al. 2019; Hemmati et al. 2019b; Masters et al. 2019; Turner et al. 2019), as well as to classify light curves of supernovae and to identify other galactic transient events (Lochner et al. 2016; Charnock & Moss 2017; Sesar et al. 2017; Carrasco-Davis et al. 2018; Hinners et al. 2018; Sooknunan et al. 2018; Aguirre et al. 2019; Muthukrishna et al. 2019a, 2019b).

Here, we investigate the capabilities of ML algorithms in finding galaxies with variable active galactic nuclei (AGNs). Powered by the accretion of matter onto supermassive black holes (SMBH) residing in the center of galaxies, AGNs shape the evolution and structure of their host galaxies through various feedback mechanisms (Bower et al. 2006; Cattaneo et al. 2006; Croton et al. 2006; Sijacki et al. 2007; Hopkins et al. 2012; Dubois et al. 2013). Measuring the number-density and rate of occurance of AGNs therefore enables us to study the formation and growth of SMBHs and their host galaxies across cosmic time (Peterson 1997). Due to different "feeding mechanisms," AGNs exhibit variations in brightness over a range of wavelengths on timescales ranging from minutes to years (Fitch et al. 1967). Variations on short timescales are likely caused by disk instabilities (Kawaguchi et al. 1998), while variations on longer timescales are dominated by the fueling of gas into the nuclear regions and regulation through feedback processes (e.g., Hopkins et al. 2012). The study of AGN variability levels therefore adds another dimension and allows us to learn about the internal processes such as inflows and outflows and the size of accretion disks around SMBHs (Shields 1978).

In this Letter, we demonstrate the capabilities of self-organizing maps (SOM; Kohonen 1982, 1990), an unsupervised ML algorithm, in identifying AGNs displaying long-term brightness variability. This algorithm has been widely used in the past, for example to study radio galaxies (Torniainen et al. 2008; Ralph et al. 2019), variable stars (Brett et al. 2004; Armstrong et al. 2016), and exoplanet transit curves (Armstrong et al. 2017) as well as to derive photometric redshifts (Carrasco Kind & Brunner 2014; Masters et al. 2015; Hemmati et al. 2019a, 2019b) and to classify gravitational waves (Rampone et al. 2013). We apply the SOM algorithm to a set of nonparametric variability estimators, which allows us to maintain domain knowledge of the data properties encapsulated in these estimators (e.g., noise, sampling rate, and selection function), and offers the flexibility of learning-based nonlinear classifications that can optimally combine these estimators for classification. Such techniques will be especially valuable for extrapolating knowledge from the deep and well-sampled parts of future surveys, such as LSST, Euclid, and *WFIRST*, to the wide and shallow parts with poor sampling. We emphasize that our work serves as a proof of concept, and the methods described here can be further refined and extended (e.g., to flux variability at multiple wavelengths).

The *WISE* AGN sample, light curves, and different variability estimators are presented in Section 2. In Section 3, we apply the SOM algorithm to our sample and compare its performance to a deep-learning regression fitting method. We conclude in Section 4. Magnitudes are expressed in the AB system (Oke & Gunn 1983) unless stated otherwise. We use a standard ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$, and $\Omega_\Lambda = 0.7$.

2. *WISE* AGN Sample

2.1. Sample Selection and Light Curves

The AGNs used in this study are color selected from a sample of 14,000 AGNs in the 270 deg^2 Sloan Digital Sky Survey (SDSS) Stripe 82 field (Jiang et al. 2014). The details are outlined in Prakash et al. (2019), a summary is provided in the following.

The color selection follows the criteria of Stern et al. (2012), using the $W1$ and $W2$ filters of the *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) centered on 3.4 μm and 4.6 μm. Specifically, we apply a cut at $W1 - W2 > 0.8\,\mathrm{mag_{Vega}}$ and restrict ourselves to $W2 > 15\,\mathrm{mag_{Vega}}$. About 50% of the candidates have confirmed redshifts and lie between $0.1 < z < 2.2$ with a median of $z \sim 0.5$.

Once the AGN candidates are identified, their time-sampled photometry is measured on *WISE* $W1$ single exposure (Level 1b) images in apertures of 6″. Bad pixels indicated by the bad pixel mask are excluded. Several standard stars are used to correct the photometry for aperture losses. The light curves are generated using only high-quality frames well separated from the South Atlantic Anomaly and bright moon light by selecting `qual_frame` = 10, `SAA_SEP` > 0, and `MOON_SEP` > 24 as suggested by the *WISE* team.[3]

The final light curves include all data from *WISE* and *NEOWISE* (Mainzer et al. 2011) over the past 10 yr. Note that a ∼3.5 yr gap around MJD 55,725 arises during the hibernation period of the telescope. *WISE* observed a single patch of sky multiple times during each visit, leading to multiple observations within typically ∼1–2 days. Since our focus here is on long-term variability, we combine these observations using median statistics, which is robust against photometric outliers. The uncertainty on the combined flux of a single visit is estimated via the weighted average

$$\sigma_\mathrm{tot}^2 = \frac{1}{\sum_{i=1}^{k}\frac{1}{\sigma_i^2}}, \tag{1}$$

where σ_i are the corresponding uncertainties on the k measured fluxes. This resampling makes the long-term variability of the light curves more apparent while increasing the signal-to-noise. For the purpose of testing ML methods, we only use AGNs whose light curves have five or more $\geqslant 5\sigma$ measurements. Although this significantly reduces the sample of AGNs, it makes the variability detection and measurement more robust. Furthermore, we focus only on the $W1$ filter as it is deeper and better time-sampled than $W2$. Our final sample consists of 8309 AGNs. To obtain a training sample, we subsequently classify the light curve of these AGNs visually in "nonvariable" (7558) and "variable" (751). In addition, we split the last category into monotonically increasing (66) and decreasing (98) light curves, while the remaining 587 vary irregularly (Figure 1).

2.2. Definition of Parametric and Nonparametric Estimators of Variability

Before applying ML methods, we define a set of estimators to characterize the variability. We distinguish between parametric and nonparametric estimators.

Parametric estimators are derived using the Gaussian process (GP) framework in Python (`GPy`[4]). We use a `RBF` Gaussian kernel and run 100 iterations of optimization on each light curve using a `L-BFGS-B` optimizer, which is typically sufficient for the parameters of the best-fit models to converge. The resulting fit is characterized by a variance and a length-scale parameter. The former is equivalent to a variability estimator, while the latter describes the timescale of a period. As nonparametric estimators we use the χ^2 test, standard deviation (σ_w), median absolute deviation (MAD), interquartile range (IQR), robust median statistics (RoMS), normalized excess variance (σ_NXS^2), peak-to-peak variability (ν), and the inverse von Neumann ratio ($1/\eta$). These estimators are described in detail in Sokolovsky et al. (2017) and we refer to their paper for exact definitions.

All estimators are computed for each of the 8309 light curves in our sample. The computation of the parametric estimators takes about 30 minutes of CPU time on a 3.1 GHz processor and the results depend on the random initial conditions for each fit in some single cases (∼5%). Nonparametric estimators are more advantageous for real-time classification as their computation requires only seconds. Furthermore, their measurement is repeatable.

3. Finding Variable AGNs

In the following, we identify variable AGNs using the unsupervised SOM algorithm implemented in the Python library `mvpa2`[5] (Hanke et al. 2009). Subsequently, we compare its performance to a more commonly used supervised deep-learning multilayer neural network algorithm that is part of the Python *TensorFlow* package.[6]

3.1. Metrics for Performance Evaluation

To compare the performance of different ML algorithms as well as the impact of different estimators, we use a common metric known as the confusion matrix, which we here define in its normalized form as

$$\mathcal{C} = \frac{1}{T}\begin{pmatrix} \mathrm{TN} & \mathrm{FP} \\ \mathrm{FN} & \mathrm{TP} \end{pmatrix}, \tag{2}$$

where TN, FP, FN, and TP denote true-negative, false-positive, false-negative, and true-positive, respectively, and T is the total sample size (TN+TP+FN+FP). From this, we derive standard metrics such as purity ($\mathcal{P}$) and completeness ($\mathcal{R}$),[7]

$$\mathcal{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \text{ and } \mathcal{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{3}$$

as well as the accuracy

$$\mathrm{ACC} = \mathrm{diag}(\mathcal{C}) = \frac{\mathrm{TP} + \mathrm{TN}}{T}, \tag{4}$$

[3] http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec2_4b.html

[4] https://sheffieldml.github.io/GPy/

[5] http://www.pymvpa.org

[6] http://www.tensorflow.org/

[7] Note that purity and completeness are equivalent to precision and recall.

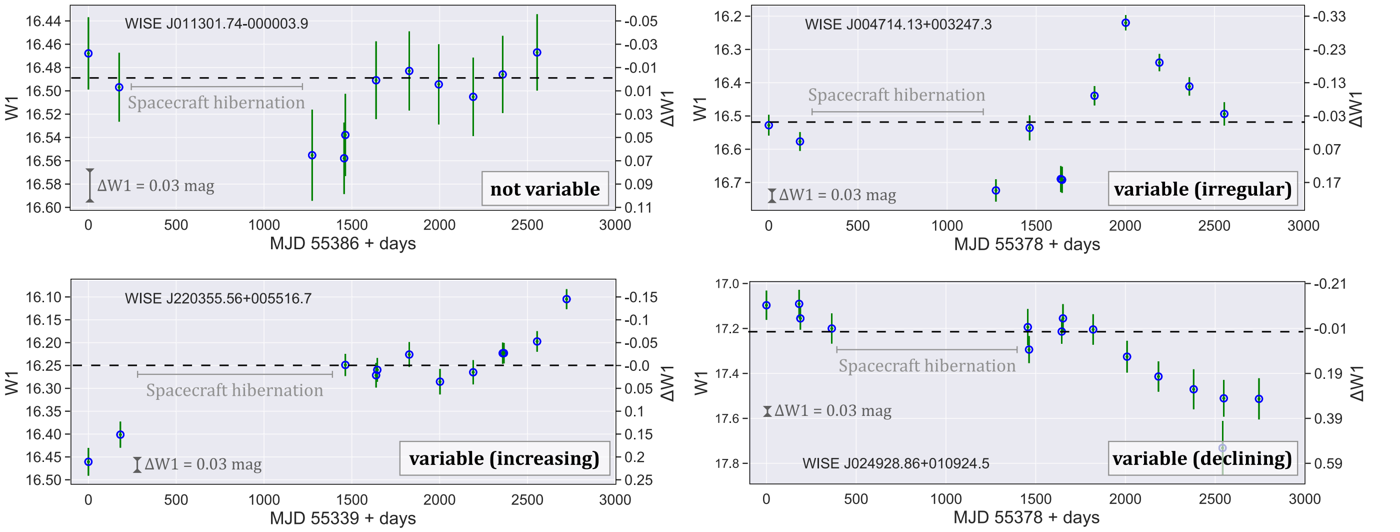

Figure 1. Representative examples of *WISE W*1 (3.4 μm) light curves in the four visual categories. The photometric uncertainties ($\leqslant$0.03 mag) are indicated for a sense of scale and the dashed line shows the median. Note the different scales of the y-axis in the plots.

the Matthews correlation coefficient (MCC, Matthews 1975)[8]

$$\mathrm{MCC} = \frac{(\mathrm{TP}\cdot\mathrm{TN}) - (\mathrm{FP}\cdot\mathrm{FN})}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}}, \tag{5}$$

and the F_1 score[9] is defined as

$$F_1 = 2\cdot\left(\frac{\mathcal{P}\cdot\mathcal{R}}{\mathcal{P}+\mathcal{R}}\right). \tag{6}$$

The MCC has several advantages over F_1 and is generally preferred for assessing the performance of a classification algorithm. For example, MCC does not depend on which outcomes are classified as positive or negative and takes correctly into account TN and FN events.

3.2. Classification with SOM

The SOM algorithm reduces an N-dimensional data set (composed of N estimators or parameters) to a two-dimensional grid of $m \times n$ cells. The algorithm preserves topological information as distances in this two-dimensional space map directly to distances in N dimensions. This makes the SOM a powerful tool for visualizing correlations in high-dimensional data sets. In detail, the SOM algorithm is initialized by the number of iterations (I), as well as a length-scale parameter (λ), learning rate[10] (L_i), and radius factor (σ_i). The latter two are decreased with iterations i, in the `mvpa2` implementation of the SOM by the factor $e^{-i/\lambda}$. In the following, we choose as initial values $\sigma_0 = \max(m, n)$ and $L_0 = 0.05$, as well as $\lambda = I/\sigma_0$. For a more detailed review of the algorithm, see, e.g., Masters et al. (2015).

3.2.1. Simulations

We first test the SOM on simulated light curves. For this, we create flat (nonvariable) as well as sinusoidal light curves with varying frequency and phase. These curves are perturbed to achieve similar noise properties as the real photometry and we also apply a time sampling similar to that of real observations. The 7700 simulated curves include 10% variable light curves, reflecting the visually derived fraction in our flux-limited *WISE* AGN sample.

We calculate all the estimators outlined in Section 2.2 and normalize and rescale them to their median and a range between 0 and 1, respectively. To train the SOM, we choose a random subsample containing 80% of the total sample (the training sample). We adopt a SOM-size of 30 × 30 cells and run 200 iteration with an initial learning rate of $L_0 = 0.05$. We test different values for the latter two (50 iterations and learning rates between 0.005 and 0.5) and find changes in the performance of less than 1%. The number of cells is chosen to optimize the performance of the algorithm. Specifically, fewer cells result in a coarser classification, hence a less clear separation of variable and nonvariable light curves. On the other hand, more cells decrease the number of light curves per cell and result in a nonuniform coverage of the map within a decrease in performance. Overall, we find these choices to be optimal in our case.

Figure 2 shows the fraction of variable light curves in each SOM-cell. The cyan contours encompass cells with a variable fraction of more than 50%. The SOM algorithm automatically groups variable light curves around the cells at (8, 7) while nonvariable light curves are distributed at larger distances. We can then quantify the performance of the algorithm by mapping the test sample (the other 20% of the total sample) onto the map. The mapping happens instantaneously for this sample size, which is a strength of the SOM algorithm. Through this mapping, each test light curve gets assigned to an SOM-cell and is then classified as variable if more than 50% of the light curves from the training sample in that cell are variable. This choice of fraction maximizes the metrics and is therefore used throughout this work. In the following, we assume this binary classification is variable/nonvariable, but note that it is possible

[8] MCC is defined between −1 and +1 with −1 (+1) indicating perfect disagreement (agreement) and 0 meaning the algorithm performs as well as random guessing.

[9] F_1 is defined between 0 and 1.

[10] The learning rate determines how fast the model is updated per iteration. Commonly, the learning rate is decreased over time for convergence.

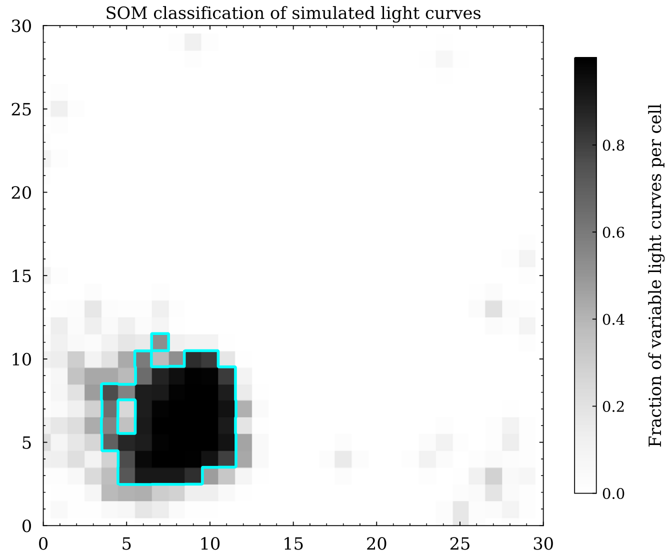

Figure 2. Test of our algorithm on simulated light curves. Shown is the fraction of truly variable light curves per SOM-cell (the cyan contour encompasses cells with a fraction higher than 50%). We are able to identify variable light curves with a purity of 91% and a completeness of 79%.

to derive a continuous scoring output for each test AGN, which would allow us to establish a confidence for each classification. Using the metrics introduced above, we quantify the success of identifying variable light curves in the test sample as

$$\mathcal{C}_{\mathrm{simulation}} = \begin{pmatrix} 0.87 & 0.01 \\ 0.03 & 0.10 \end{pmatrix}, \tag{7}$$

with ACC = 0.97, MCC = 0.83, and an F_1 score of 0.85. The purity and completeness of the classification are 91% and 79%, respectively, suggesting a "contamination" of nonvariable light curves of 9% in a variable sample selected by our algorithm. Note that the SOMs are randomly initialized, hence these numbers can change for different representations. By running the algorithm multiple times, we find that these changes are on the order of ±0.01 (or 1% in per-cent notation).

3.2.2. Application to Observed Light Curves

Having shown that the SOM is a powerful tool to identify variability, we now apply the algorithm to real light curves. For this, we normalize and rescale the estimators measured for our *WISE* AGN sample as described above. A training fraction of 80% (6647 AGN) is again used to train the SOM. To generate a smooth SOM map, we remove 100 AGNs from the training sample for which at least one estimator lies in the top 1% of the distribution. We find that this clipping improves the performance of the algorithm slightly. Note that this cut is not applied to the test sample. Based on our simulations, we adopt an SOM-size of 30 × 30 cells and run 200 iterations at a learning rate of 0.05.

Figure 3 shows the fraction of variable AGNs per SOM-cell for the data (the cyan contour encompasses cells with a variable fraction >50%). For educational purposes, we indicate the location of the AGN shown in Figure 1 and list on the right light curves contained in the cells at (20, 18), (22, 12), (15, 7), and (13, 12). Cells 1 and 2 are dominated by variable light curves while cells 3 and 4 contain predominantly nonvariable AGNs. In this specific representation of the SOM, variable light curves cluster around the cell at (21, 18). For our basic SOM classification, we find

$$\mathcal{C}_{\mathrm{SOM}} = \begin{pmatrix} 0.85 & 0.01 \\ 0.04 & 0.09 \end{pmatrix}, \tag{8}$$

with ACC = 0.94, MCC = 0.72, an F_1 score of 0.75 and a purity (completeness) of 86% (66%). While the SOM can easily identify variable light curves, we find that splitting into subcategories of variability (i.e., irregular, decreasing, and increasing) cannot be achieved robustly. This is not surprising given their small relative number compared to the total sample (164 out of 8309).

By mapping the training sample back onto the SOM cells and computing the median of each estimator per cell, we can visualize and study the correlations of estimators with variability. These elements of the *Kohonen* layer are shown in Figure 4 (panels (1a)–(1h)). The light red contours show cells with a variable fraction >50% (see cyan contours in Figure 3). We also show the distribution of the parametric estimators (variance and length scale) on the map (panels (2a) and (2b)). Most of the estimators correlate well with the fraction of variable AGNs per cell. Notably, the χ^2, RoMS, and $1/\eta$ estimators correlate best with variability as they peak around the cells with the highest fraction of variable AGNs. Note that the inverse von Neumann ratio ($1/\eta$) is the only estimator discussed here that takes into account the correlation between two successive data points in a time series. Specifically, $1/\eta$ is large for smoothly varying curves, such as smoothly decreasing or increasing light curves. On the other hand, the ratio is small for fluctuations on short timescales (as in highly variable AGNs or nonvariable AGNs with large photometric uncertainties). The other estimators show a wider extent on the maps, suggesting less correlation with variability. This is likely due to degeneracies in the low signal-to-noise regime. Specifically, the estimators MAD, IQR, and σ^2_{NXS} are offset to the northeast and show high values also for nonvariable AGNs. We note that the same behavior is seen on the Kohonen maps of the simulated light curves. Such degeneracies may arise because the MAD and IQR estimators do not take into account the photometric uncertainties. As a consequence, a truly nonvariable light curve, poorly sampled in time, can mimic changing brightness (hence a high MAD and IQR) solely due to large photometric errors. Indeed, the average signal-to-noise of the observations is lower in these regions. A similar explanation holds for σ^2_{NXS}. One could think of removing these estimators for the training of the SOM to improve the identification of variable light curves. We investigate this by training the algorithm only on the χ^2, RoMS, and $1/\eta$ estimators. However, it turns out that overall the performance is slightly worse, suggesting that the removed estimators contain some important information for the classification. Specifically, we find

$$\mathcal{C}_{\mathrm{SOM}}|_{\chi^2,\mathrm{RoMS},1/\eta} = \begin{pmatrix} 0.86 & 0.02 \\ 0.04 & 0.08 \end{pmatrix}, \tag{9}$$

with ACC = 0.94, MCC = 0.71, an F_1 score of 0.73, a purity (completeness) of 84% (65%).

The variance parametric estimator shows a good correlation with variability in contrast to the length-scale estimator, which displays significant scatter and no clear relation. The latter is anticorrelated with the variance estimator as expected—it correctly identifies variable AGNs with a short length scale (i.e., period); however, the opposite is not true. Including the variance estimator to train the SOM results in a similar

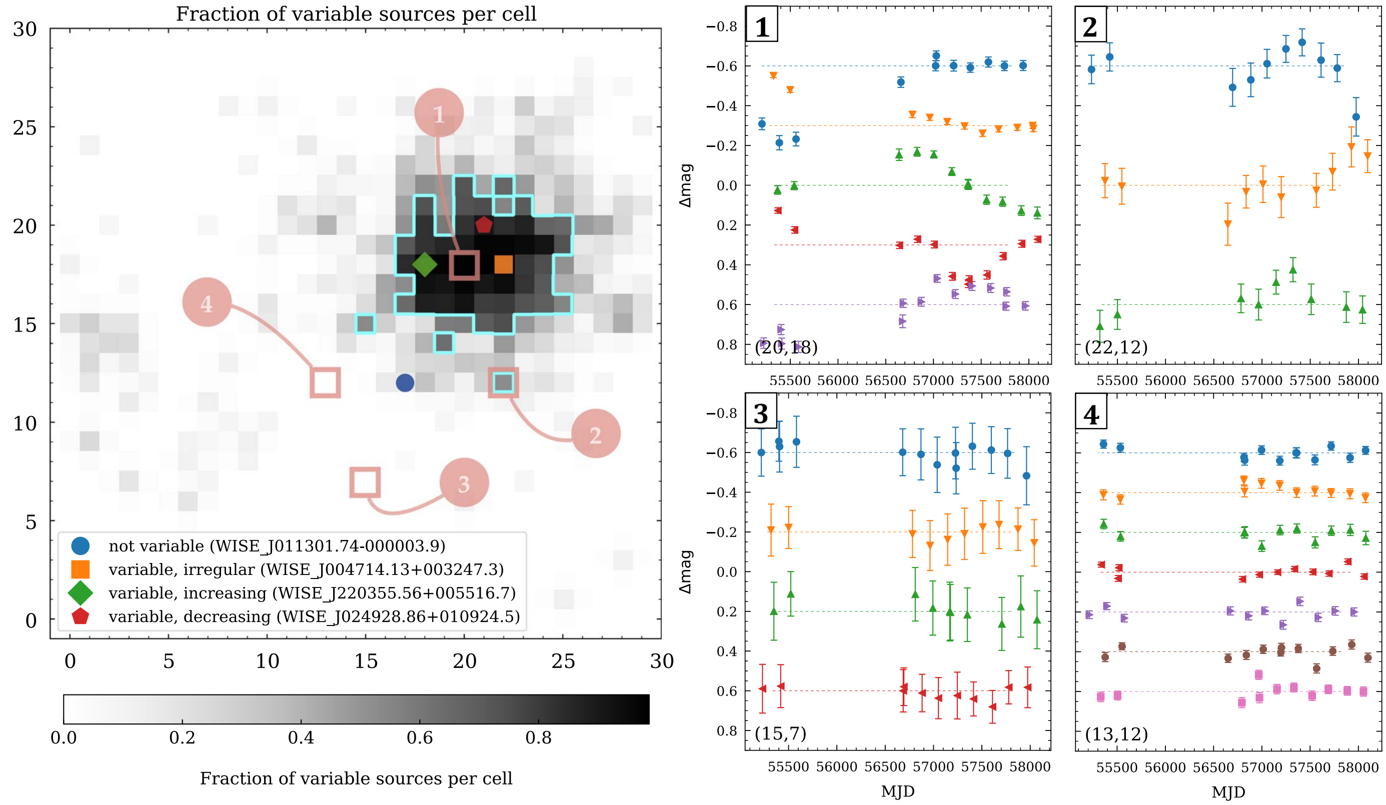

Figure 3. Fraction of observed variable AGN light curves per SOM-cell (the cyan contour encompasses cells >50%). The SOM algorithm classifies an AGN as variable with a purity of 86% and completeness of 66%. The color symbols indicate the location of the AGN shown in Figure 1. The panels on the right show the light curves (offset by a constant factor) residing in each of the four SOM-cells indicated by the boxes. Cells 1 and 2 are dominated by variable AGNs, while cells 3 and 4 contain mostly nonvariable light curves. All light curves are plotted on the same scale (normalized to median).

performance (ACC = 0.94, MCC = 0.70, an F_1 score of 0.73, a purity of 84%, and completeness of 64%), hence the benefit of including this parametric estimator is questionable, also given that its computation requires two orders of magnitude more CPU time compared to the computation of the nonparametric estimators. In addition, we test if the performance can be increased by down-sampling the training sample to an equal number of variable and nonvariable AGNs (the latter are randomly selected). Indeed, we achieve a higher purity and completeness (93% and 90%) determined on the training sample. However, the performance is worse if determined on the full sample. This is likely because of the small size of the training sample (1502 AGNs, out of which 751 are variable). Giving a different weighting to the classes of AGNs could improve the performance in future analyses, but implementing this is beyond the scope of this paper.

3.3. Comparison with Deep-learning Neural Networks

Finally, we compare the performance of the SOM with a more commonly used supervised deep-learning neural network approach—here the Multilayer Perceptrons (MLP) method implemented in the Python *TensorFlow* package. We build a sequential model with three Dense layers. Two of them with 64 and 128 nodes and a rectified linear unit (`tf.nn.relu`) activation and one with two nodes (yes/no) and normalized exponential (`tf.nn.softmax`) activation. The model is compiled using a stochastic gradient descent (SGD) optimizer with a sparse categorical cross entropy. The deep-learning algorithm is trained on the training sample using the same set of nonparametric estimators as used to train the SOM. We find a confusion matrix of

$$\mathcal{C}_{\text{deep}} = \begin{pmatrix} 0.89 & 0.01 \\ 0.04 & 0.05 \end{pmatrix}, \tag{10}$$

with ACC = 0.94, MCC = 0.65, and an F_1 score of 0.67. The purity and completeness are 79% and 58%, respectively, comparable to the SOM algorithm. With a similar amount of training time as needed to train the SOM, we find a comparable performance between the two methods. Compared to other optimizers (e.g., the *adams* or *RMSprop* optimizer) we find that the SGD optimizer shows the best performance. We also test a convolutional neural network method and find very similar results.

4. Conclusions

In this Letter, we demonstrate the combination of domain knowledge of how to measure variability with the flexibility and optimization that ML-based approaches bring to large data sets. Using simulated light curves of different variability, we demonstrate how unsupervised self-organizing maps can be used to identify variable AGN light curves in a heterogeneous data set. This provides powerful means of using ML to identify variability in the presence of photometric noise, selection functions, and heterogeneous sampling in future surveys. We apply our method to a sample of 8309 AGN light curves, out of which ~10% are identified as variable by our visual inspection. The SOM algorithm can recover variable AGNs with a purity of 86% and completeness of 66%. The training of the SOM

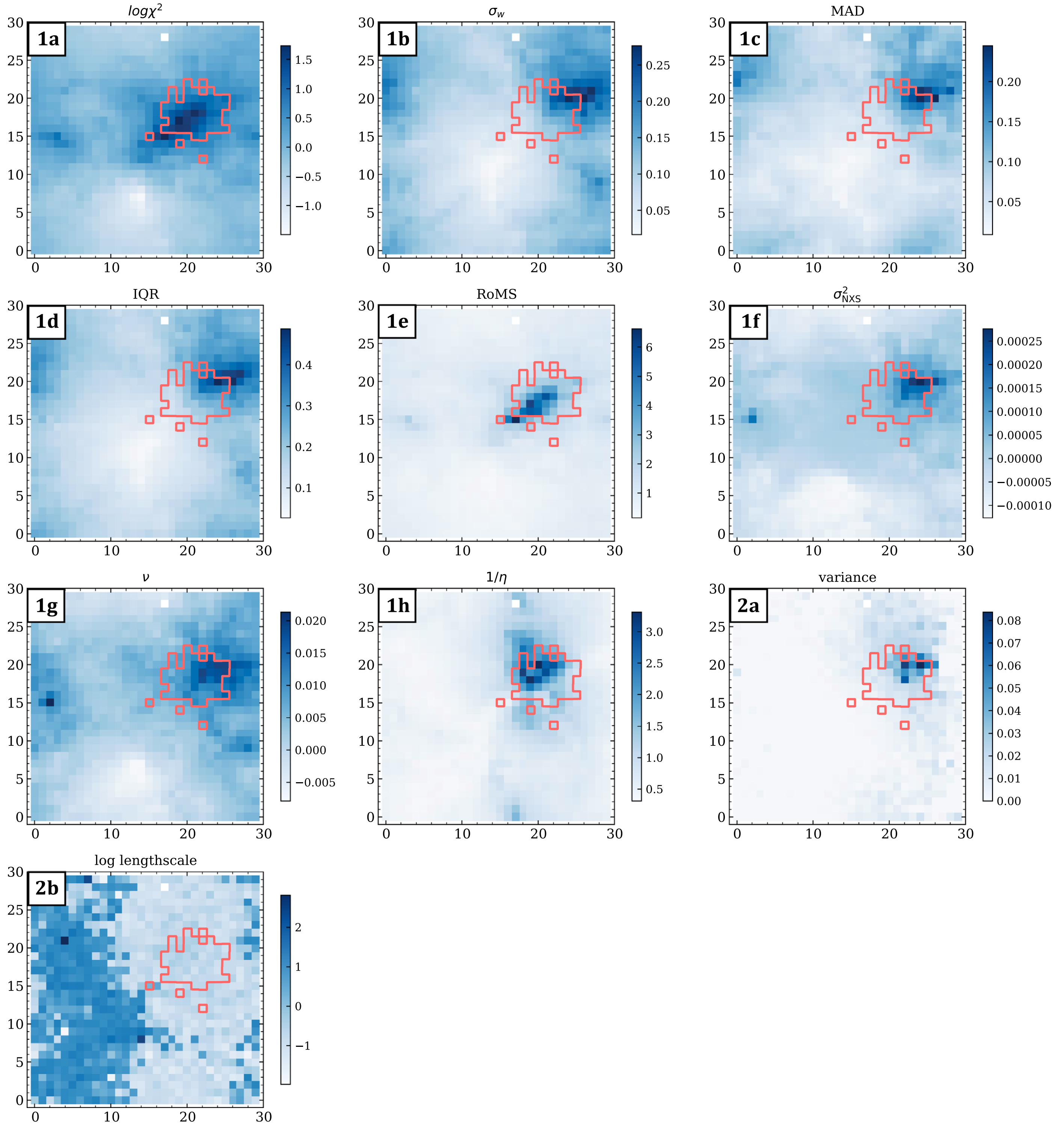

Figure 4. Distribution of different estimators on the 30 × 30 cells SOM map. The light red contours show cells with a variable fraction of >50%. Only nonparametric estimators (panels (1a)−(1h)) are used to train the SOM. Most of the estimators correlate well with variability. The estimators σ_w, MAD, IQR, and σ^2_{NXS} show offsets indicative of degeneracies in the low signal-to-noise limit.

(done only once) takes less than 100 s (∼4000 objects, 8 estimators) on a 3.1 GHz processor. The classification of a test sample of similar size is instantaneous and can be achieved in real-time for much larger data sets. In the same CPU time and identical test situations, supervised deep-learning networks perform comparable to the SOM but lack the visualization of the correlation between estimators and the “fitted” quantity and cannot easily be applied to data sets with missing labels (i.e., unsupervised). The SOM framework is powerful to reveal and study connections between variability and other physical properties and processes (e.g., connection between variability and properties of the host galaxy or accretion models). We here used variable AGN as use-case, but our method can be applied to any light curves to identify supernovae, transiting exoplanets, pulsars, and other transient events in large data sets.

We thank the anonymous referee for the useful comments that improved this paper. We also thank J. Tan for proofreading the manuscript. This work was supported by Joint Survey Processing (JSP) at IPAC/Caltech, which is aimed at combined analysis of Euclid, LSST, and *WFIRST*. This work was supported by Joint Survey Processing (JSP) at IPAC/Caltech which was funded by NASA grant 80NM0018F0803. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

ORCID iDs

Andreas L. Faisst https://orcid.org/0000-0002-9382-9832
Abhishek Prakash https://orcid.org/0000-0003-4451-4444
Peter L. Capak https://orcid.org/0000-0003-3578-6843
Bomee Lee https://orcid.org/0000-0003-1954-5046

References

Aguirre, C., Pichara, K., & Becker, I. 2019, MNRAS, 482, 5078
Armstrong, D. J., Kirk, J., Lam, K. W. F., et al. 2016, MNRAS, 456, 2260
Armstrong, D. J., Pollacco, D., & Santerne, A. 2017, MNRAS, 465, 2634
Bonjean, V., Aghanim, N., Salomé, P., et al. 2019, A&A, 622, A137
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Brett, D. R., West, R. G., & Wheatley, P. J. 2004, MNRAS, 353, 369
Carrasco Kind, M., & Brunner, R. J. 2014, MNRAS, 438, 3409
Carrasco-Davis, R., Cabrera-Vives, G., Förster, F., et al. 2018, arXiv:1807.03869
Cattaneo, A., Dekel, A., Devriendt, J., Guiderdoni, B., & Blaizot, J. 2006, MNRAS, 370, 1651
Charnock, T., & Moss, A. 2017, ApJL, 837, L28
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Davidzon, I., Laigle, C., Capak, P. L., et al. 2019, arXiv:1905.13233
Dubois, Y., Gavazzi, R., Peirani, S., & Silk, J. 2013, MNRAS, 433, 3297
Fitch, W. S., Pacholczyk, A. G., & Weymann, R. J. 1967, ApJL, 150, L67
Hanke, M., Halchenko, Y. O., Sederberg, P. B., et al. 2009, Neuroinformatics, 7, 37
Hemmati, S., Capak, P., Masters, D., et al. 2019a, ApJ, 877, 117
Hemmati, S., Capak, P., Pourrahmani, M., et al. 2019b, arXiv:1905.10379
Hinners, T. A., Tat, K., & Thorp, R. 2018, AJ, 156, 7
Hopkins, P. F., Hayward, C. C., Narayanan, D., & Hernquist, L. 2012, MNRAS, 420, 320
Jiang, L., Fan, X., Bian, F., et al. 2014, ApJS, 213, 12
Kawaguchi, T., Mineshige, S., Umemura, M., & Turner, E. L. 1998, ApJ, 504, 671
Kohonen, T. 1982, Biological Cybernetics, 43, 59
Kohonen, T. 1990, IEEE, 78, 1464
Krakowski, T., Małek, K., Bilicki, M., et al. 2016, A&A, 596, A39
Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. 2016, ApJS, 225, 31
Mainzer, A., Bauer, J., Grav, T., et al. 2011, ApJ, 731, 53
Masters, D., Capak, P., Stern, D., et al. 2015, ApJ, 813, 53
Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2019, ApJ, 877, 81
Matthews, B. 1975, Biochimica et Biophysica Acta, Protein Structure, 405, 442
Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019a, arXiv:1904.00014
Muthukrishna, D., Parkinson, D., & Tucker, B. 2019b, arXiv:1903.02557
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Peterson, B. M. 1997, An Introduction to Active Galactic Nuclei (Cambridge: Cambridge Univ. Press)
Prakash, K., et al. 2019, ApJ, submitted
Ralph, N. O., Norris, R. P., Fang, G., et al. 2019, arXiv:1906.02864
Rampone, S., Pierro, V., Troiano, L., & Pinto, I. M. 2013, IJMPC, 24, 1350084
Sesar, B., Hernitschek, N., Mitrović, S., et al. 2017, AJ, 153, 204
Shields, G. A. 1978, BAAS, 10, 690
Sijacki, D., Springel, V., Di Matteo, T., & Hernquist, L. 2007, MNRAS, 380, 877
Siudek, M., Małek, K., Pollo, A., et al. 2018, A&A, 617, A70
Sokolovsky, K. V., Gavras, P., Karampelas, A., et al. 2017, MNRAS, 464, 274
Sooknunan, K., Lochner, M., Bassett, B. A., et al. 2018, arXiv:1811.08446
Speagle, J. S., & Eisenstein, D. J. 2017a, MNRAS, 469, 1186
Speagle, J. S., & Eisenstein, D. J. 2017b, MNRAS, 469, 1205
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
Torniainen, I., Tornikoski, M., Turunen, M., et al. 2008, A&A, 482, 483
Turner, S., Kelvin, L. S., Baldry, I. K., et al. 2019, MNRAS, 482, 126
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868