Permutation Entropy and Statistical Complexity Analysis of MODIS Satellite NDVI Data for Discerning *Xylella fastidiosa* Infection in Olive Groves in Southern Italy

Luciano Telesca and Rosa Lasaponara
Institute of Methodologies for Environmental Analysis
National Research Council
Tito, Italy
email: luciano.telesca@cnr.it; rosa.lasaponara@cnr.it

***Abstract*—This study utilizes the Complexity–Entropy Causality Plane (CECP) to investigate the temporal dynamics of Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) time series, focusing on olive groves in Southern Italy. The primary objective is to identify the presence of *Xylella fastidiosa*, an aggressive phytobacterium responsible for the devastating Olive Quick Decline Syndrome (OQDS). Our results demonstrate that the application of the CECP framework facilitates a highly effective discrimination between infected and healthy olive orchards. Our analysis reveals that infected datasets are characterized by a lower permutation entropy (H) and larger statistical complexity (C) than uninfected ones. By employing an embedding dimension of $d_x = 4$ a wider separation between infected and infected datasets in the CECP is obtained, evidenced also by an increased Mahalanobis distance and superior Receiver Operating Characteristic (ROC) classification performance. These findings validate the CECP as an optimal diagnostic metric for identifying subtle irregularities within MODIS NDVI data due to *Xylella fastidiosa*-induced degradation of vegetation.**

Index Terms*—MODIS; NDVI; CECP; *Xylella fastidiosa

I. Introduction

Invasive pests and non-native plant bacteria are increasingly recognized as major threats worldwide, capable of causing severe plant diseases with far-reaching effects on both natural ecosystems and agricultural production. A prominent example is *Xylella fastidiosa*, considered one of the most hazardous plant pathogens globally. This bacterium is responsible for serious diseases that particularly affect crops such as grapevines, citrus, and olive trees. Within the European Union, focusing specifically on olives, it is estimated that *Xylella fastidiosa* could lead to annual production losses of up to 5.5 billion euros, jeopardizing approximately 70% of the EU's mature olive tree output. Consequently, early detection of infections caused by this pathogen is essential to implement effective strategies for minimizing its economic and ecological impact [2].

In 2013, *Xylella fastidiosa* was first detected in southeastern Italy [3]. Since then, the pathogen has spread rapidly to several other European countries, leading to a serious phytosanitary crisis [4].

Traditionally, visual inspection has been the primary method for detecting *Xylella fastidiosa*, due to its simplicity and low cost. However, the accuracy of this approach is highly dependent on the subjective assessment of disease severity.

In recent years, Remote Sensing (RS) techniques have gained increasing attention for monitoring vegetation dynamics, including changes caused by plant pathogens [5]. Several RS-based studies in phytopathology have focused on methods leveraging multi-temporal and multi-spectral satellite imagery to track land cover changes. Statistical approaches, such as principal component analysis [6] and curve-fitting techniques [7], have proven effective in identifying vegetation alterations across large areas.

The potential of MODIS satellite data for monitoring pest and parasite outbreaks at both landscape and field scales has been explored in [8], [9], where MODIS vegetation indices proved effective in tracking the degradation of pine tree vegetation caused by the parasite *Toumeyella parvicornis*. Additionally, analyses using multifractal detrended fluctuation analysis and the Fisher-Shannon informational method applied to MODIS evapotranspiration time series have demonstrated that these satellite-derived data provide a reliable means of assessing the pathogenic effects of *Xylella fastidiosa* on olive trees [10]–[12].

The Normalized Difference Vegetation Index (NDVI) is a well-established indicator of vegetation vigor, photosynthetic activity, and canopy structure, and is particularly suitable for monitoring the effects of Xylella fastidiosa infection in olive groves. The pathogen disrupts xylem water transport, inducing water stress, leaf desiccation, and progressive canopy loss, which result in reduced near-infrared reflectance and increased red reflectance, effectively captured by NDVI. Satellite-based NDVI time series, especially those derived from the MODIS sensor, provide consistent, long-term observations with high temporal resolution, enabling large-scale monitoring of olive-growing areas. In infected groves, NDVI temporal profiles typically exhibit reduced seasonal amplitude, altered phenological patterns, and increased temporal irregularity compared to healthy areas. When combined with nonlinear time-series analysis and statistical complexity measures, MODIS NDVI can reveal early dynamic changes associated with infection, making NDVI a robust, scalable, and cost-effective tool for the surveillance and monitoring of Xylella fastidiosa impacts.

In this study, we analyze time series of MODIS NDVI

pixels covering both infected and healthy olive groves in Southern Italy using the Complexity–Entropy Causality Plane (CECP) that enables a topological characterization of time series through the distribution of ordinal patterns, providing a robust alternative to traditional information-theoretic and fluctuation-based approaches previously applied to this type of data.

The Complexity–Entropy Causality Plane (CECP) was first introduced by Rosso *et al.* in 2007 [13] and has since become a widely used diagnostic tool. The CECP represents a two-dimensional space combining normalized permutation entropy (H) with a statistical complexity measure (C). While H measures the degree of randomness or unpredictability in a time series, C captures non-trivial correlations and structural patterns. By projecting time series onto the $H \times C$ plane, it is possible to differentiate between various stochastic and deterministic processes, allowing the identification of transitions between chaotic and noisy regimes [16]. The CECP has been successfully applied in environmental science, including the analysis of air pollution dynamics [17], river discharge variability under anthropogenic influence [19], and rainfall time series [18].

In Section II, the applied methods of time series analysis will be described and in Section III the investigated dataset will be presented. Section IV will show the obtained results that will be discussed in Section V.

II. Methods

A. The permutation entropy and statistical complexity

Permutation entropy is computed by determining the Shannon entropy of a probability distribution derived from ordinal (or permutation) patterns. These patterns are extracted by applying a sliding window across a time series to generate symbolic representations. This specific distribution, often termed the *ordinal distribution*, is estimated following the symbolization framework proposed by Bandt and Pompe [1].

To formalize this methodology, consider a discrete time series $\{x_t\}_{t=1,\dots,N_x}$. The sequence is decomposed into $n_x = N_x - (d_x - 1)\tau_x$ overlapping vectors. Each vector consists of $d_x > 1$ elements sampled at a temporal distance of $\tau_x \geq 1$. For a given set of parameters, the p-th partition is defined as:

$$w_p = \left(x_p, x_{p+\tau_x}, x_{p+2\tau_x}, \dots, x_{p+(d_x-1)\tau_x}\right), \tag{1}$$

where $p = 1, \dots, n_x$ represents the partition index. The Bandt–Pompe approach relies on two fundamental parameters: the *embedding dimension* d_x and the *embedding delay* τ_x. While the seminal work by Bandt and Pompe [1] focused on consecutive observations ($\tau_x = 1$), the framework was subsequently generalized to include arbitrary delays. Because the selection of d_x and τ_x is critical to the analysis, extensive research has been dedicated to identifying optimal parameter configurations for different signal types [13] [16].

Following the partitioning of the signal, each vector w_p is mapped to a specific permutation $\pi_p = (r_0, r_1, \dots, r_{d_x-1})$. This permutation represents the unique ordering of the set $\{0, 1, \dots, d_x - 1\}$ required to arrange the components of w_p in non-decreasing order. Formally, the ordinal pattern π_p is established by the relation:

$$x_{p+r_0\tau_x} \leq x_{p+r_1\tau_x} \leq \cdots \leq x_{p+r_{d_x-1}\tau_x}. \tag{2}$$

To ensure a unique symbolic representation in the presence of ties (equal values), the initial chronological sequence is maintained. Specifically, if $x_{p+r_k\tau_x} = x_{p+r_h\tau_x}$, the condition $r_k < r_h$ is imposed for any $k < h$ [1].

To demonstrate this mapping, let us examine a sample sequence $x_t = (6, 4, 3, 3, 8, 10)$ with parameters $d_x = 4$ and $\tau_x = 1$. The initial window is defined as $w_1 = (6, 4, 3, 3)$. Sorting these values gives the sequence $3 \leq 3 < 4 < 6$, which links back to the original positions $x_{1+2}, x_{1+3}, x_{1+1}, x_{1+0}$. Consequently, the resulting ordinal pattern for this segment is $\pi_1 = (2, 3, 1, 0)$.

Although Bandt and Pompe originally mentioned that ties could be resolved by introducing a negligible amount of observational noise to the dataset, this approach is rarely adopted in empirical studies. While high-resolution data typically minimize the frequency of identical values, the choice of tie-breaking strategy becomes a critical consideration when processing low-resolution or heavily quantized signals [16].

After converting all partitions into their respective symbolic representations, we obtain a symbolic sequence $\{\pi_p\}_{p=1,\dots,n_x}$. The corresponding ordinal probability distribution, $P = \{\rho_i(\pi_i)\}_{i=1,\dots,n_\pi}$, is determined by calculating the relative frequency of each unique permutation within the sequence. Formally, this is expressed as:

$$\rho_i(\pi_i) = \frac{\text{count}(\pi_p = \pi_i)}{n_x}, \tag{3}$$

where π_i represents one of the $n_\pi = d_x!$ possible permutations of order d_x.

Using the distribution P, we calculate its Shannon entropy, defining the *permutation entropy* as:

$$SH(P) = -\sum_{i=1}^{n_\pi} \rho_i(\pi_i) \log_2 \rho_i(\pi_i). \tag{4}$$

This metric serves as a proxy for the structural stochasticity of the time series. Specifically, $SH \approx \log_2 n_\pi$ implies a nearly random process, whereas $SH \approx 0$ reflects highly predictable or deterministic behavior.

It is well-established that the estimator in (4) exhibits a negative bias, particularly for finite datasets [20], [22]. To mitigate this, we employ the Miller–Madow correction [21], which refines the estimate as follows:

$$S(P) = SH(P) + \frac{N_0 - 1}{2N}, \tag{5}$$

where N_0 denotes the count of permutations with non-zero probability and N represents the total number of observed patterns.

Finally, to facilitate comparisons across different embedding dimensions, we utilize the *normalized permutation entropy*:

$$H(P) = \frac{S(P)}{\log_2 n_\pi}, \tag{6}$$

which maps the entropy values onto the unit interval $[0, 1]$.

The embedding dimension d_x defines the symbolic resolution of the system, with the total number of potential ordinal patterns given by $n_\pi = d_x!$. To achieve a statistically consistent estimation of the ordinal probability distribution, Bandt and Pompe [1] suggest selecting d_x such that $d_x! \ll N_x$. Simultaneously, the embedding delay τ_x determines the relevant time scale for the dynamical analysis. In our case $\tau_x = 1$.

Rosso *et al.* [13] introduced the CECP as a diagnostic tool, designed to differentiate between different processes.

The CECP integrates the normalized permutation entropy H from (6) with the statistical complexity measure C to establish a bivariate diagnostic space, where C is analyzed as a function of H. Indicating as $P = \{\rho_i(\pi_i)\}_{i=1,\dots,n_\pi}$ the observed ordinal distribution and as $U = \{1/n_\pi\}_{i=1,\dots,n_\pi}$ the uniform distribution,

$$C(P) = \frac{D(P, U)\, H(P)}{D_{\max}}, \tag{7}$$

where the Jensen–Shannon divergence $D(P, U)$ is calculated as:

$$D(P, U) = S\left(\frac{P + U}{2}\right) - \frac{1}{2}S(P) - \frac{1}{2}S(U). \tag{8}$$

The term $D_{\max}$ serves as a normalization constant, ensuring that the complexity measure is bounded. It represents the maximum possible divergence between P and U, which occurs when the system is concentrated in a single state (i.e., P follows a Kronecker delta distribution, $\delta_{1,i}$). This constant is defined as:

$$D_{\max} = -\frac{1}{2}\left[(n_\pi + 1)\log_2(n_\pi + 1) - 2\log_2(2n_\pi) + \log_2 n_\pi\right]. \tag{9}$$

Unlike entropy, C vanishes at both poles of the dynamical spectrum: complete order (periodic or constant signals) and total disorder (white noise). By quantifying the interplay between information storage and structural dissimilarity, C provides essential diagnostic details that complement the randomness quantified by H, making the CECP a powerful tool for detecting subtle regime shifts in complex systems.

B. The ROC analysis

Receiver Operating Characteristic (ROC) analysis serves as a fundamental framework for assessing the diagnostic efficacy of binary classifiers [24]. In this context, instances are categorized into two discrete classes, typically labeled as positive or negative. The predictive performance is determined by the resulting confusion matrix, which consists of four possible outcomes: True Positive (TP), where a positive instance is accurately identified; False Negative (FN), where a positive instance is erroneously classified as negative; True Negative (TN), signifying the correct identification of a negative instance; and False Positive (FP), where a negative instance is incorrectly assigned a positive label [24].

To quantify classification accuracy, we define the *True Positive rate* (TPr) and the *False Positive rate* (FPr) as follows:

$$\mathrm{TPr} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{10}$$

$$\mathrm{FPr} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{11}$$

The ROC curve characterizes the trade-off between the TPr (sensitivity) and the FPr (1 − specificity) across a range of decision thresholds. To construct the curve for a specific parameter, the data from classes C_1 and C_2 are consolidated and ranked. A sliding threshold T is then varied across the data range. Assuming the distribution mean of C_1 exceeds that of C_2, a TP is recorded when a value from C_1 is $\geq T$, while a FP occurs when a value from C_2 is $\geq T$. The optimal classification threshold is identified as the point on the ROC curve with the minimum Euclidean distance to the ideal coordinates $(0, 1)$, representing the highest achievable discriminative power.

C. The Mahalanobis distance

The statistical divergence between two datasets can be quantified using the Mahalanobis distance D_M, defined as:

$$D_M(\bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) = \sqrt{(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)^\top \mathbf{S}_p^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)}, \tag{12}$$

where the indices i and j represent the respective groups under comparison, and $\bar{\mathbf{x}}$ denotes the vector of mean values for the parameters of interest. The term $\mathbf{S}_p$ refers to the pooled covariance matrix. Consequently, D_M provides a robust multivariate measure of the dissimilarity between groups, accounting for the internal geometry of the data distribution.

III. Data and study area

The data exhibit a spatial resolution of 250m and an 16-day sampling rate. They are readily accessible online [27] and are also present in the Google Earth Engine (GEE) cloud database.

The NDVI is obtained from the visible (red) and near-infrared (NIR) by using the following formula

$$NDVI = \frac{NIR - red}{NIR + red}. \tag{13}$$

The differential reflectance in these bands provides a means of monitoring density and vigor of green vegetation growth using the spectral reflectivity of solar radiation. Green leaves commonly have larger reflectances in the near-infrared than in the visible range. As the leaves undergo water stress, become diseased or die, they become more yellow and reflect significantly less in the near-infrared range. Therefore the NDVI is regarded as a reliable indicator for land cover variations [14] since its temporal evolution is strongly linked to changes in the state of surface.

In this study, we analyzed three datasets of MODIS NDVI pixels derived from olive groves affected by *Xylella fastidiosa* corresponding to sites where the infection was first reported in 2015 (X2015), in 2016 (X2016), and in 2017 (X2017). As a reference, we also examined datasets of pixels from two healthy olive groves located in the Foggia and Matera areas, which exhibit climatic and topographic conditions almost similar to those of the infected sites. Before performing the CECP analysis of the data, we firstly removed from the datasets outliers. Outliers in the NDVI time series were detected using a moving median filter. For a given series, a sliding window of size 5 was used

to compute the local median, and points that deviated from this local median by more than three times the local median were removed. This approach is robust to short-term fluctuations and noise, such as cloud contamination or sensor errors, because it evaluates each point relative to its local temporal context rather than to the global statistics of the series. The NDVI time series span the period from 2010 to 2022, with each pixel represented by 288 samples. Given the length of the available NDVI time series ($N = 288$ samples per pixel), the analysis was restricted to embedding dimensions $d_x \in \{3, 4\}$, following the conservative criterion proposed by Amigó *et al.* [15], which requires $5 \cdot d_x! \leq N$ to ensure a statistically reliable estimation of the ordinal probability distribution. This condition is satisfied for $d_x = 3$ ($5 \cdot 3! = 30 \leq 288$) and $d_x = 4$ ($5 \cdot 4! = 120 \leq 288$), but violated for $d_x = 5$ ($5 \cdot 5! = 600 > 288$). We processed only data with a gap percentage below 25%. Figure 1 shows the time variation of the spatial means of NDVI data from each of the five investigated sites.

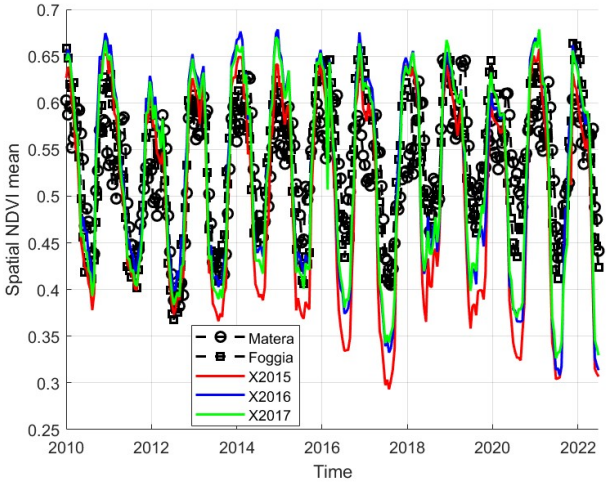

Figure 1. Temporal evolution of the spatial means of NDVI for the investigated sites.

IV. Results

Figure 2 reports the box-plot distributions of H and C for $d_x = 3$. The analysis reveals clear differences in the dynamical behavior of the uninfected datasets (Matera and Foggia) compared to the infected datasets (X2015, X2016, and X2017). In particular, the uninfected datasets are characterized by a larger median entropy with respect to the X2015–X2017 sites, indicating a lower degree of intrinsic organization in the underlying dynamics. In contrast, the infected datasets exhibit a larger decrease in H, consistent with diminished stochasticity in the signal evolution.

An opposite tendency is observed for the complexity measure. The uninfected datasets show lower median values of C. Conversely, the X2015–X2017 datasets display larger complexity levels. The simultaneous decrease in entropy and increase in complexity observed in the latter datasets suggests a transition from disordered, noise-dominated toward more structured states.

The pairwise analysis conducted in the CECP (Figure 3) further quantifies the separation between the datasets. The CECP plots reveal Matera and Foggia datasets consistently occupying the low-complexity/high-entropy region of the plane; while the X-series datasets cluster significantly toward the high-complexity/low-entropy region. The Mahalanobis distance for Matera vs X-data range from $D_M = 1.53$ to $D_M = 0.95$; while a bit smaller Mahalanobis distance is for Foggia vs X-data, ranging from $D_M = 1.13$ to $D_M = 0.56$. These values confirm a quantifiable divergence in the statistical quantifiers of the uninfected and the infected sites. The ROC curves (Figure 4) yield significant Area Under the Curve (AUC) values from 0.86 to 0.74 for Matera vs X-data and from 0.79 to 0.66 for Foggia vs X-data, providing a measure of diagnostic accuracy. The discrimination between uninfected and infected (X2015 and X2016) datasets is good (but larger for Matera). However, for the more recent X2017 dataset, the AUC indicates a less efficient discrimination between the two types of datasets.

Similar results (not shown) are obtained for $d_x = 4$. Tables I to IV report the values of D_M, thresholds, AUC, TPr, and FPr for H and C for $d_x = 3$ and $d_x = 4$. From a synoptical evaluation of the reported results, $d_x = 4$ provides a bit more refined characterization of the signal dynamics, offering higher sensitivity in discriminating between uninfected and infected series compared to $d_x = 3$.

Table I. SUMMARY OF PERMUTATION ENTROPY (H) AND COMPLEXITY (C) METRICS FOR MATERA (dx = 3).

Metric	Matera vs X2015	Matera vs X2016	Matera vs X2017
D_M	1.530	1.265	0.953
Permutation Entropy (H)			
Threshold$_H$	0.969	0.971	0.973
AUC$_H$	0.859	0.816	0.745
FPR$_H$	0.259	0.315	0.419
TPR$_H$	0.856	0.829	0.797
Complexity (C)			
Threshold$_C$	0.031	0.031	0.031
AUC$_C$	0.858	0.814	0.743
FPR$_C$	0.181	0.178	0.187
TPR$_C$	0.778	0.692	0.568

Table II. SUMMARY OF PERMUTATION ENTROPY (H) AND COMPLEXITY (C) METRICS FOR FOGGIA (dx = 3).

Metric	Foggia vs X2015	Foggia vs X2016	Foggia vs X2017
D_M	1.127	0.836	0.561
Permutation Entropy (H)			
Threshold$_H$	0.972	0.972	0.980
AUC$_H$	0.791	0.734	0.665
FPR$_H$	0.218	0.299	0.290
TPR$_H$	0.677	0.681	0.549
Complexity (C)			
Threshold$_C$	0.029	0.030	0.024
AUC$_C$	0.789	0.731	0.662
FPR$_C$	0.349	0.338	0.450
TPR$_C$	0.809	0.718	0.707

V. Conclusion and Future work

The results obtained in this study provide a robust dynamical framework for characterizing the impact of *Xylella fastidiosa* on

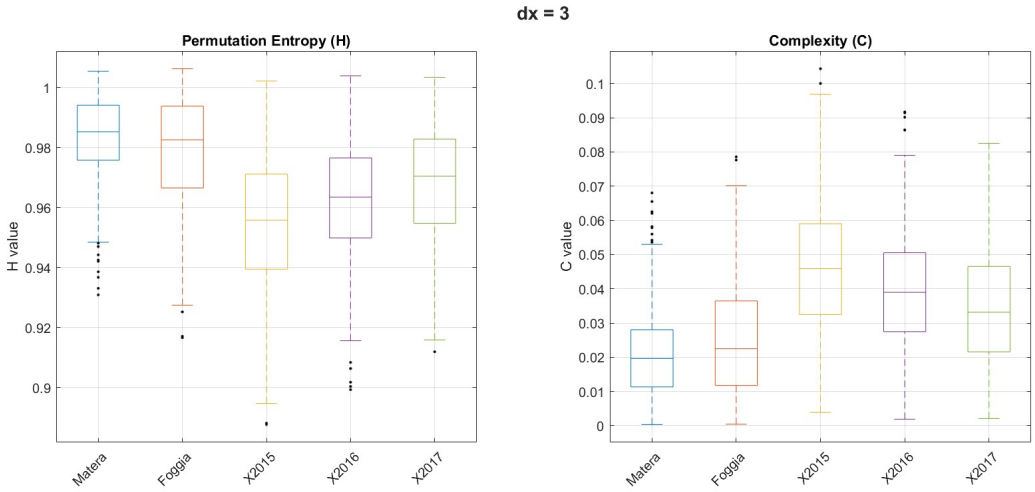

Figure 2. Box-plots of H and C for the investigated datasets and embedding dimension $d_x = 3$.

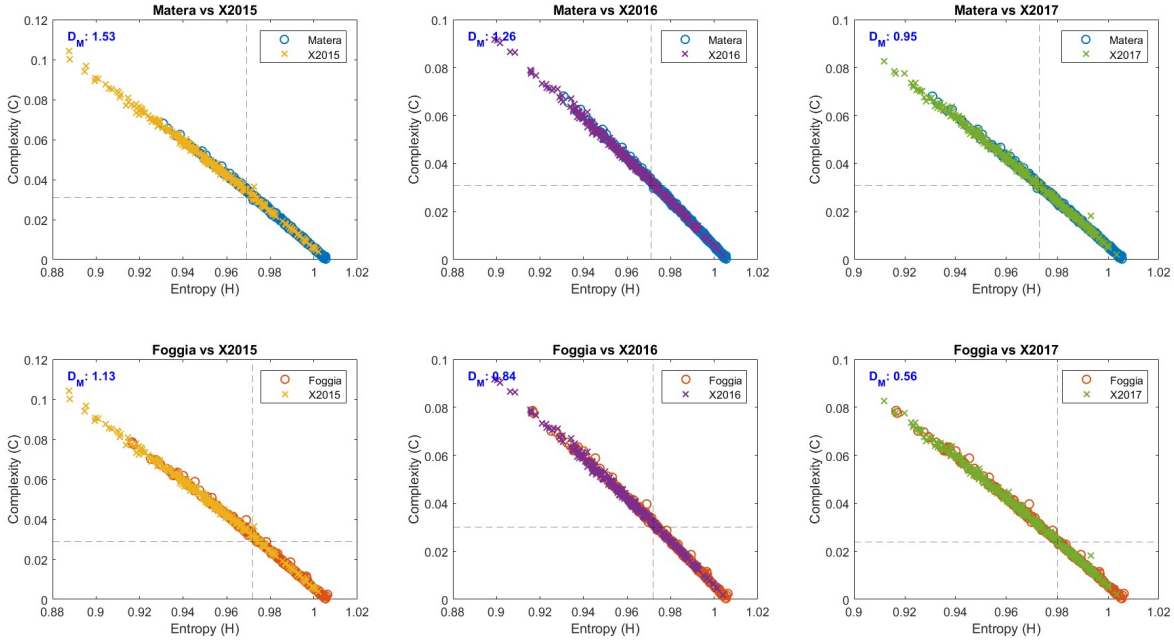

Figure 3. CECPs for $d_x = 3$. The dotted vertical and horizontal lines mark the optimal thresholds for H and C.

Table III. SUMMARY OF PERMUTATION ENTROPY (H) AND COMPLEXITY (C) METRICS FOR MATERA (dx = 4).

Metric	Matera vs X2015	Matera vs X2016	Matera vs X2017
D_M	1.666	1.382	1.111
Permutation Entropy (H)			
Threshold$_H$	0.936	0.939	0.943
AUC$_H$	0.878	0.834	0.777
FPR$_H$	0.175	0.242	0.297
TPR$_H$	0.806	0.785	0.747
Complexity (C)			
Threshold$_C$	0.090	0.089	0.090
AUC$_C$	0.872	0.827	0.772
FPR$_C$	0.208	0.216	0.208
TPR$_C$	0.818	0.754	0.647

Table IV. SUMMARY OF PERMUTATION ENTROPY (H) AND COMPLEXITY (C) METRICS FOR FOGGIA (dx = 4).

Metric	Foggia vs X2015	Foggia vs X2016	Foggia vs X2017
D_M	1.135	0.837	0.617
Permutation Entropy (H)			
Threshold$_H$	0.928	0.931	0.931
AUC$_H$	0.789	0.729	0.673
FPR$_H$	0.243	0.325	0.421
TPR$_H$	0.718	0.694	0.694
Complexity (C)			
Threshold$_C$	0.102	0.086	0.096
AUC$_C$	0.780	0.720	0.665
FPR$_C$	0.267	0.423	0.315
TPR$_C$	0.725	0.780	0.588

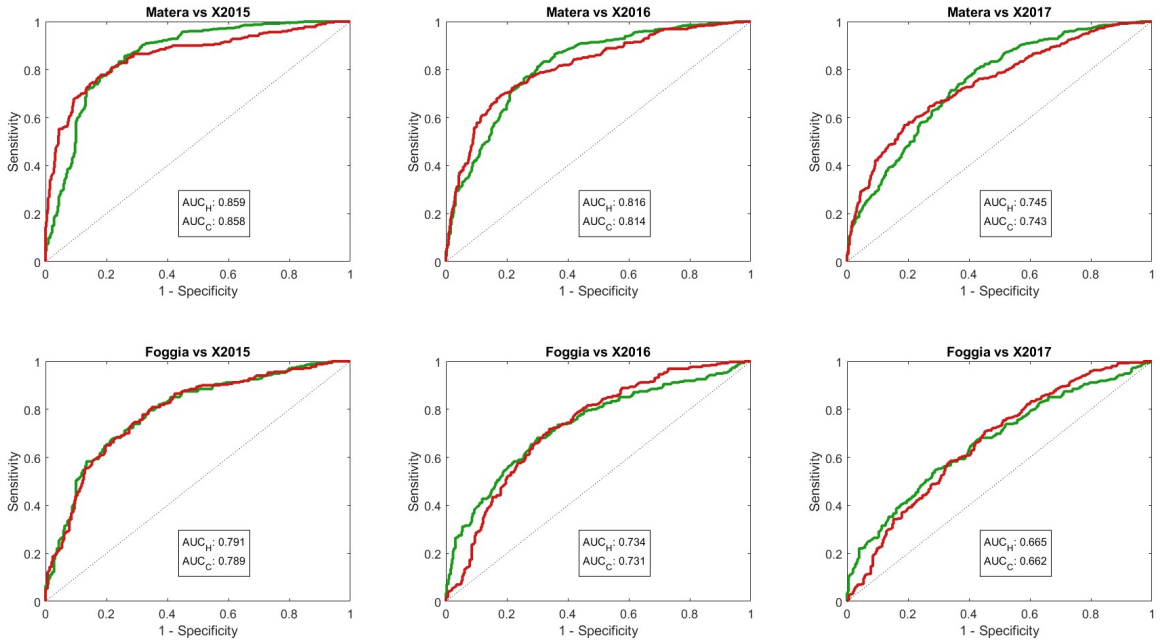

Figure 4. ROC curves (bottom panels) for $d_x = 3$. The green and red ROC curves refer to H and C respectively.

olive groves.

The transition from uninfected sites (Matera and Foggia) to the infected datasets (X2015–X2017) signifies a profound shift in the system's dynamical regime: the high values of H and low levels of C observed in Matera and Foggia indicate a system in a flexible, resilient equilibrium; the signal fluctuates freely, reflecting a "natural noise" where a healthy canopy absorbs environmental variations without losing its dynamical "plasticity". In contrast, the simultaneous decrease in H and increase in C in the X-datasets suggests that the infection imposes a specific "structure" onto the signal; the resulting desiccation is not a random process but a structured, ordered degenerative pathway. This is also observed in Figure 1, where the NDVI time series of the infected pixels reveals a progressive structured decay—most notably following the 2015 outbreak—with the seasonal troughs becoming progressively deeper. In fact, in the infected sites, the thinning of the olive tree canopy causes the 250 m pixel to be increasingly influenced by the underlying herbaceous vegetation. Unlike the stable, evergreen nature of healthy olive trees, grass follows a significantly more extreme seasonal cycle, characterized by rapid growth followed by quick senescence. The dominance of these rigid seasonal cycles makes the signal dynamics in the infected sites more "predictable" compared to the stochastic stability of a healthy evergreen canopy. Consequently, the infected sites exhibit a decrease in permutation entropy, as the biological signal is replaced by a more deterministic, seasonally-driven pattern. Furthermore, the transition from a stable perennial state to a structured decline introduces non-trivial temporal correlations that enhance the C.

The values of the AUC and the Mahalanobis Distances confirm that non-linear analysis is a powerful diagnostic tool. Matera serves as the "gold standard" for olive grove purity, making the contrast with infected sites highly evident. The D_M from 1.53 to 0.95 quantifies this clear divergence. A smaller D_M between 1.13 and 0.56 and slightly lower AUC values suggest that Foggia, while healthy, may possess a dynamics closer to the infected Apulian sites. This could be attributed to shared pedoclimatic conditions or similar land management practices; nevertheless, the statistical distinction remains significant.

In summary, the approach based on the CECP does more than merely identify the presence of *Xylella fastidiosa*; it quantifies the pathogen's ability to dismantle the plant's biological regulation mechanisms. This diagnostic framework effectively characterizes the transition where the inherent flexibility of a healthy life state is replaced by the rigid, structured dynamics of a pathological decline in olive trees.

References

[1] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Physical Review Letters*, vol. 88, p. 174102, 2002.

[2] European Commission Joint Research Centre, "Monitoring the impacts of *Xylella fastidiosa*," https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/monitoring-impacts-xylella-fastidiosa-2019-11-12_en, [retrieved: May, 2025].

[3] M. Saponari, D. Boscia, F. Nigro, and G. P. Martelli, "Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy)," *Journal of Plant Pathology*, vol. 95, pp. 659–668, 2013.

[4] M. Jeger, J. Bragard, D. Caffier, T. Candresse, E. Chatzivassiliou, K. Dehnen-Schmutz, and G. Lovell, "Scientific opinion on the updated pest categorisation of *Xylella fastidiosa*," *EFSA Journal*, vol. 16, p. 5357, 2018.

[5] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, pp. 265–274, 2009.

[6] J. M. Bioucas-Dias and M. A. Figueiredo, "A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, pp. 2992–3004, 2007.

[7] G. Chen and D. Needell, "Compressed sensing and dictionary learning," *Finite Frame Theory: A Complete Introduction to Overcompleteness*, vol. 73, pp. 201–241, 2016.

[8] L. Telesca, M. Lovallo, A. Pisarenko, and R. Lasaponara, "Exploring long-term anomalies in the vegetation cover of peri-urban parks using the Fisher–Shannon method," *Entropy*, vol. 24, p. 1784, 2022.

[9] L. Telesca, M. Lovallo, A. Pisarenko, and R. Lasaponara, "Urban and peri-urban vegetation monitoring using satellite MODIS NDVI time series, singular spectrum analysis, and Fisher–Shannon statistical method," *Sustainability*, vol. 15, p. 11039, 2023.

[10] L. Telesca, N. Abate, F. Faridani, M. Lovallo, and R. Lasaponara, "Revealing traits of phytopathogenic status induced by *Xylella fastidiosa* in olive trees by analyzing multifractal and informational patterns of MODIS satellite evapotranspiration data," *Physica A*, vol. 629, p. 129163, 2023.

[11] L. Telesca, N. Abate, F. Faridani, M. Lovallo, and R. Lasaponara, "Discerning *Xylella fastidiosa*-infected olive orchards in the time series of MODIS Terra satellite evapotranspiration data by using the Fisher–Shannon analysis and the multifractal detrended fluctuation analysis," *Fractal and Fractional*, vol. 7, p. 466, 2023.

[12] L. Telesca, N. Abate, M. Lovallo, and R. Lasaponara, "Investigating the impact of *Xylella fastidiosa* on olive trees by the analysis of MODIS Terra satellite evapotranspiration time series by using the Fisher information measure and the Shannon entropy: A case study in Southern Italy," *Remote Sensing*, vol. 16, p. 1242, 2024.

[13] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes, "Distinguishing noise from chaos," *Physical Review Letters*, vol. 99, p. 154102, 2007.

[14] K. E. Huemmrich, T. A. Black, P. G. Jarvis, J. H. McCaughey, and E. G. Hall, "High temporal resolution NDVI phenology from micrometeorological radiation sensor," *Journal of Geophysical Research*, vol. 104, p. 27935, 1999.

[15] J. M. Amigó, M. B. Kennel, and L. Kocarev, "The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical systems," *Physica D: Nonlinear Phenomena*, vol. 228, no. 1, pp. 77–85, 2007.

[16] L. Zunino, M. C. Soriano, and O. A. Rosso, "Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach," *Physical Review E*, vol. 86, p. 046210, 2012.

[17] F. Ma, Q. Fan, and G. Ling, "Complexity-entropy causality plane analysis of air pollution series," *Fluctuation and Noise Letters*, vol. 21, p. 2250011, 2021.

[18] A. S. A. D. Silva, I. D. D. C. Barreto, M. Cunha-Filho, R. S. C. Menezes, B. Stosic, and T. Stosic, "Multiscale complexity analysis of rainfall in Northeast Brazil," *Water*, vol. 13, p. 3213, 2021.

[19] M. Suriano, L. F. Caram, and O. A. Rosso, "Daily streamflow of Argentine rivers analysis using information theory quantifiers," *Entropy*, vol. 26, p. 56, 2024.

[20] L. Paninski, "Estimation of entropy and mutual information," *Neural Computation*, vol. 15, pp. 1191–1253, 2003.

[21] G. A. Miller, "Note on the bias of information estimates," in *Information Theory in Psychology: Problems and Methods*, H. Quastler, Ed. Glencoe, IL: Free Press, pp. 95–100, 1955.

[22] J. De Gregorio, D. Sánchez, and R. Toral, "Entropy estimators for Markovian sequences: A comparative analysis," *Entropy*, vol. 26, p. 79, 2024.

[23] R. López-Ruiz, H. L. Mancini, and X. Calbet, "A statistical measure of complexity," *Physics Letters A*, vol. 209, pp. 321–326, 1995.

[24] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.

[25] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series," *Physical Review E*, vol. 80, p. 046103, 2009.

[26] M. Saponari, A. Giampetruzzi, G. Loconsole, D. Boscia, and P. Saldarelli, "*Xylella fastidiosa* in olive in Apulia: Where we stand," *Phytopathology*, vol. 109, pp. 175–186, 2019.

[27] NASA Land Processes Distributed Active Archive Center, "LP DAAC – Land Products," https://lpdaac.usgs.gov, [retrieved: May, 2025].

[28] L. Laipelt, R. H. B. Kayser, A. S. Fleischmann, A. Ruhoff, W. Bastiaanssen, T. A. Erickson, and F. Melton, "Long-term monitoring of evapotranspiration using the SEBAL algorithm and Google Earth Engine cloud computing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 81–96, 2021.

[29] MathWorks File Exchange, "Fast natural visibility graph (NVG) for MATLAB," https://www.mathworks.com/matlabcentral/fileexchange/70432, [retrieved: May, 2025].

[30] G. Bounova and O. de Weck, "Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles," *Physical Review E*, vol. 85, p. 016117, 2012.

[31] B. A. Gonçalves, L. Carpi, O. A. Rosso, and M. G. Ravetti, "Time series characterization via horizontal visibility graph and information theory," *Physica A*, vol. 464, pp. 93–102, 2016.

[32] K. Schneider, W. van der Werf, M. Cendoya, M. Mourits, J. A. Navas-Cortés, A. Vicent, and A. Oude Lansink, "Impact of *Xylella fastidiosa* subspecies *pauca* in European olives," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, pp. 9250–9259, 2020.