



HAL
open science

Time-lagged recurrence: A data-driven method to estimate the predictability of dynamical systems

Chenyu Dong, Davide Faranda, Adriano Gualandi, Valerio Lucarini, Gianmarco Mengaldo

► **To cite this version:**

Chenyu Dong, Davide Faranda, Adriano Gualandi, Valerio Lucarini, Gianmarco Mengaldo. Time-lagged recurrence: A data-driven method to estimate the predictability of dynamical systems. Proceedings of the National Academy of Sciences of the United States of America, 2025, 122 (20), <10.1073/pnas.2420252122>. <hal-04709054v2>

HAL Id: hal-04709054

<https://hal.science/hal-04709054v2>

Submitted on 17 May 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Time-lagged recurrence: A data-driven method to estimate the predictability of dynamical systems

Chenyu Dong^a, Davide Faranda^{b,c,d}, Adriano Gualandri^{e,f}, Valerio Lucarini^g, and Gianmarco Mengaldo^{a,h,1}

Affiliations are included on p. 9.

Edited by Manlio De Domenico, Università degli Studi di Padova, Padova, Italy; received October 3, 2024; accepted March 20, 2025, by Editorial Board Member Alan Hastings

Nonlinear dynamical systems are ubiquitous in nature and they are hard to forecast. Not only they may be sensitive to small perturbations in their initial conditions, but they are often composed of processes acting at multiple scales. Classical approaches based on the Lyapunov spectrum rely on the knowledge of the dynamic forward operator, or of a data-derived approximation of it. This operator is typically unknown, or the data are too noisy to derive its faithful representation. Here, we propose a data-driven approach to analyze the local predictability of dynamical systems. This method, based on the concept of recurrence, is closely linked to the well-established framework of local dynamical indices. When applied to both idealized systems and real-world datasets arising from large-scale atmospheric fields, our approach proves its effectiveness in estimating local predictability. Additionally, we discuss its relationship with other local dynamical indices, and how it reveals the scale-dependent nature of predictability. Furthermore, we explore its link to information theory, its extension that includes a weighting strategy, and its real-time application. We believe these aspects collectively demonstrate its potential as a powerful diagnostic tool for complex systems.

predictability | dynamical systems | data-driven methods | local dynamical indices

Systems that evolve in time according to specific rules are known as dynamical systems, and the associated field of study is called dynamical system theory. There is a rich history on this topic, starting from the 1600s, when Newton developed calculus and classical mechanics (1), to the 1800s, when Poincaré developed geometrical approaches to the study of dynamical systems (2). In the 1900s, the field flourished, with the pioneering contributions of Birkhoff, Kolmogorov, Moser, Arnold, Lorenz, Ruelle, Takens, May, and Feigenbaum, among others, who laid the foundations of modern dynamical system theory, including works on oscillators, chaos, and fractals, well summarized by Strogatz (3).

Many real-world systems are made of multiple interacting parts, giving rise to complex dynamical systems. A paradigm for describing such systems relies on considering high-dimensional deterministic models, which are often characterized by chaotic behavior and multiscale dynamics (4, 5). A separate paradigm relies on considering stochastic dynamics, where randomness is explicitly introduced in the system (6, 7). When constructing reduced order models of high dimensional systems, stochastic dynamics is an emergent property. This is at the core of the Hasselmann's program for climate science (8, 9) and can be rigorously justified thanks to the Mori–Zwanzig theory (10, 11). The latter also lays the foundation for the use of data-driven methods aimed at estimating the role of the unresolved scales of motion on the resolved ones (12). And, indeed, low-dimensional stochastic models have proved very valuable tools for understanding the statistical and dynamical properties of complex systems, e.g. in the area of Earth system science (13–18).

Either way, a well-known fact is that we currently observe a finite predictability horizon for systems of interest like, for example, the atmosphere (19, 20). This means that, when forecasting the future behavior of a dynamical system, the error that we make tends to increase the longer into the future we want to forecast (21).

To understand predictability and the methods commonly used to assess it, we look at the evolution of the system in the so-called phase space, that is the space of all the variables used to describe it. In the phase space usually we can identify a subset, namely a finite-size manifold of trajectories followed by the dynamical system over time and denoted by $\mathbf{x}(t)$, where \mathbf{x} is the vector of observables that depends on time t . This finite-size manifold is also known as attractor [random attractor in the case of stochastic dynamical systems (22)].

Significance

In many complex systems, predictability can be substantially state-dependent. We propose here a purely data-driven approach for estimating the local predictability at different time scales. The effectiveness of our approach is validated against existing knowledge, and its relationship to commonly used metrics in the dynamical systems literature is discussed. Our approach provides useful and convincing insights for both simple and real-world systems, including large-scale atmospheric circulation patterns in the Euro-Atlantic sector, clarifying key aspects of their predictability. Additionally, it can be used as a diagnostic tool in existing data-driven workflows to better understand the behavior of dynamical systems, opening a wide range of research opportunities.

Author contributions: C.D., D.F., and G.M. designed research; C.D., A.G., and G.M. performed research; C.D., D.F., A.G., V.L., and G.M. analyzed data; and C.D., D.F., A.G., V.L., and G.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. M.D.D. is a guest editor invited by the Editorial Board.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

¹To whom correspondence may be addressed. Email: mpegim@nus.edu.sg.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2420252122/-/DCSupplemental>.

Published May 16, 2025.

The origin of this practical unpredictability may stem either from an intrinsic stochastic nature of the system or from the ignorance of the current state of the system and/or of the dynamic rule controlling its evolution (23), or from a combination of both. Furthermore, it is well known that even low-dimensional nonlinear systems may be difficult to forecast if they exhibit an exponential divergence of close-by states (19). In such cases, the unpredictability comes from our lack of precise knowledge of the current state of the system: If we knew exactly the initial conditions, we would always get the same future evolution. A common quantity used to characterize the predictability of a system is the maximum Lyapunov exponent λ_{\max} , which measures the fastest possible rate of exponential divergence between two nearby trajectories in the phase space for an infinitely long time horizon. If we look at the evolution of two trajectories starting from two points close to each other, namely $\mathbf{x}(t_0)$ and $\mathbf{x}(t_0) + \delta\mathbf{x}$, these will diverge exponentially fast as time progresses, until their distance will plateau around a value bounded by the diameter of the attractor (24). Even more instructive is to look at the evolution of a ball around the state $\mathbf{x}(t)$. From this, one can define a full Lyapunov spectrum, made of Lyapunov exponents equal in number to the variables used to describe the phase space. The Lyapunov exponents are mean logarithmic growth rates of the lengths of the principal axes of this ball evolving under the action of the dynamic rule (4, 5, 24, 25). The existence of at least one positive Lyapunov exponent is the hallmark of chaotic dynamical systems. Therefore, measuring predictability in dynamical systems can be reframed as computing the divergence of a system's nearby trajectories over time.

It is important to highlight that the Lyapunov exponents are averaged quantities and the average is typically performed over a long trajectory. As a consequence, they serve as key indicators of the system's global predictability (4, 25), where the term global refers to average properties of the system, hence to the average predictability of the system. While the average predictability of the system can be useful, theoretical curiosity as well as operational requirements and priorities indicate the need to investigate a local notion of predictability. Indeed, the stability properties of a dynamical system can be dramatically state-dependent (26). We then wish to understand the system's predictability in a specific region of the phase space (we will refer to these local characteristics also as instantaneous, since they refer to a specific time) (27, 28). This notion has extreme relevance in the context of weather prediction (29) and key implications for devising efficient and accurate data assimilation methods (30).

To this end, local (or finite-time) Lyapunov exponents (28, 31–33) are useful quantities. However, their reliance on linearized dynamics prevents the characterization of the system outside the linear regime of perturbations. Acknowledging this limitation, finite-size Lyapunov exponents (34, 35) and nonlinear local Lyapunov exponents (36, 37) were subsequently proposed to quantify local predictability in the nonlinear regime, as they consider finite perturbations. Nonetheless, it is noteworthy that they both assume an exponential law for error growth. Although this is consistent with classic Lyapunov exponents within the linear regime, this assumption often does not hold in the nonlinear regime (i.e., for finite errors). These local indices have been applied successfully to various dynamical systems of varied complexity to gain insights into their predictability (32, 34, 36, 38), despite the aforementioned limitations.

Besides Lyapunov exponents and their variants, information theory also offers methods to quantify predictability (23, 39). Within this framework, various entropy-based indicators are usually employed as predictability measures for model-based

prediction, with varying degrees of success (40–43). However, because these methods are model-based, they are unsuitable for observational datasets. Other indices from nonlinear time series analysis also use predictability measures such as sample entropy, permutation entropy, and others (44–46); nevertheless, they are only effective with time series or systems with relatively low dimensionality. To the best of the authors' knowledge, no information theory method can infer local predictability information from high-dimensional observational data.

From the brief review of relevant literature just outlined, the above mentioned indices have several drawbacks. More specifically, Lyapunov exponents and their variants may not work well for nonlinear regimes, while information-theory-based methods may not be directly applicable to high-dimensional dynamical systems. In the next section, we leverage the recent framework proposed by Lucarini et al. (47), which uses recurrences to characterize local dynamical properties, to develop a local predictability index that differs from existing metrics, such as the Lyapunov exponents and their variants.

A Local Data-Driven Approach

Lucarini et al., (47) introduced a novel framework to measure the local properties of dynamical systems. In particular, they combine extreme value theory (48) and the Poincaré recurrence theorem (2) to construct purely data-driven local dynamical indices, that can measure the local dimension of dynamical systems and their persistence at any point in time. Under the ergodic assumption, the Poincaré recurrence theorem guarantees that the system will visit an arbitrarily small neighborhood of any state again and again as time passes. In this framework, for one given state of interest $\zeta = \mathbf{x}(t = t_\zeta)$ in the phase space, its neighboring states, also known as recurrences, are used to characterize its local dynamical properties. This can be formalized by defining the negative logarithmic distance between ζ and all other states

$$g(\zeta, \mathbf{x}(t)) = -\log[\text{dist}(\mathbf{x}(t), \zeta)], \quad [1]$$

where the dist function can be any distance metric, and by subsequently defining a quantity called exceedance for the neighboring states (i.e., recurrences):

$$\mathbf{u}(\zeta) = g(\zeta, \mathbf{x}(t)) - s(q, \zeta), \quad \forall g(\zeta, \mathbf{x}(t)) > s(q, \zeta). \quad [2]$$

In Eq. 2, $s(q, \zeta)$ is a high threshold corresponding to the q -quantile of the series $g(\zeta, \mathbf{x}(t))$, noting that having $g(\zeta, \mathbf{x}(t))$ above the threshold $s(q, \zeta)$ is a synonym of requiring a trajectory to fall within a neighborhood of ζ (defined as a ball).

Building on the above-defined quantities and leveraging extreme value theory, this framework introduces two local dynamical indices: The local dimension d and persistence Θ (see *Materials and Methods* for a detailed explanation and derivation), both of which are frequently associated with the predictability of dynamical systems (49–51). More specifically, high local dimension and low persistence are seen as an indication that the system is less predictable. The rationale for this is as follows. If a system requires more degrees of freedom (high d) to be described at a given point in the phase space, then its complexity is greater at that given point, making it possibly less predictable. Similarly, if a system leaves a neighborhood quickly (low Θ), then the dynamics in that neighborhood are fast and nonpersistent, an aspect that may indicate reduced predictability. While these arguments on predictability are qualitatively valid, and frequently correct, they approach predictability from a complexity (d), and

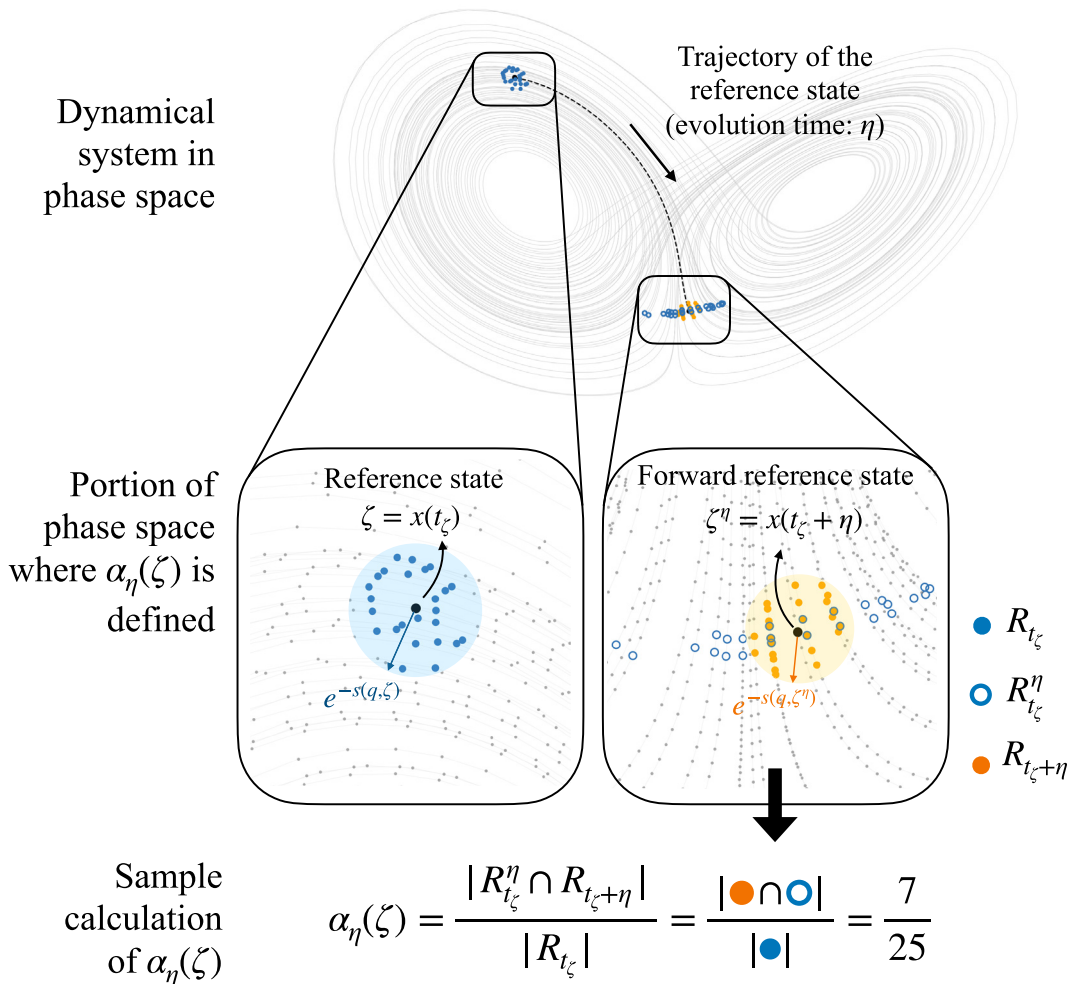


Fig. 1. Schematic illustration of the computation of $\alpha_\eta(t)$, demonstrated in the phase space of the Lorenz-63 system. This figure presents all steps involved in computing the *time-lagged recurrence* for the reference state ζ at a forecasting horizon η , namely $\alpha_\eta(\zeta)$. The panels in the second row provide a zoomed-in view of the phase space region where $\alpha_\eta(\zeta)$ is defined. Recurrences (R_{t_ζ}), forward recurrences ($R_{t_\zeta}^\eta$), and forward-reference-state recurrences ($R_{t_\zeta+\eta}$) are represented by solid blue dots, empty blue dots, and orange dots, respectively. The blue circle with radius $e^{-s(q,\zeta)}$ indicates the hypersphere used to define the neighborhood of the reference state, while the orange circle with radius $e^{-s(q,\zeta^\eta)}$ corresponds to the forward-reference-state.

mean residence time (Θ) perspective, without directly addressing predictability.

In this work, we address the predictability of dynamical systems by building upon the framework of local dynamical indices just outlined. In particular, we introduce a new quantity named “time-lagged recurrence” (also referred to as TLR) and denoted by α_η . This new quantity uses neighboring states to monitor an ensemble of trajectories, and provides an indication in a statistical sense of the local predictability of dynamical systems. This is accomplished in a purely data-driven fashion, i.e. without requiring a model nor further simplifications on the data.

The computation of α_η also relies on the concept of recurrences and the quantities defined above in Eqs. 1 and 2, with a schematic illustration presented in Fig. 1. More specifically, we identify the set of states—i.e., recurrences— R_{t_ζ} in the neighborhood of the reference state ζ , that is

$$R_{t_\zeta} = \{\mathbf{x}(t_{u_i}) : u_i(\zeta) > 0 \wedge |t_\zeta - t_{u_i}| > w\}_{i=1}^{N_\eta}, \quad [3]$$

where $u_i(\zeta)$ represents the exceedances, defined in Eq. 2, and with positive values indicating proximity to the reference state in phase space, t_ζ is the time that identifies the reference state ζ ,

N_η is the number of recurrences (i.e., the number of temporally nonadjacent neighbors) at time t_ζ , t_{u_i} is the time corresponding to the exceedance u_i , and w is the Theiler window (52) that excludes temporal neighbors of ζ . We note that if w is chosen too small that cannot exclude temporal neighbors it will lead to spurious high values of α_η .

After having identified the recurrences R_{t_ζ} , to characterize the local predictability of ζ , we first need to estimate how neighboring states evolve after time η has passed from t_ζ . To this end, we first evolve the states within R_{t_ζ} forward by time η (where η is also referred to as forecasting horizon), which leads to a set of new states, that we name forward recurrences

$$R_{t_\zeta}^\eta = \{\mathbf{x}(t + \eta) \mid \forall \mathbf{x}(t) \in R_{t_\zeta}\}. \quad [4]$$

These new states, that were evolved forward in time by η from time t_ζ , are represented by the circles with blue contours in Fig. 1, and are denoted by using the superscript $(\cdot)^\eta$.

Once the forward recurrences have been defined (Eq. 4), we need to define the forward reference state, that is, the reference state $\zeta^\eta = \mathbf{x}(t_\zeta + \eta)$, and denoted by $\zeta^\eta = \mathbf{x}(t_\zeta + \eta)$. This forward reference state will have its own exceedances $\mathbf{u}(\zeta^\eta)$ made

of N_n elements $\{u_i(\zeta^\eta)\}_{i=1}^{N_n}$, and we can compute the forward-reference-state recurrences

$$R_{t_\zeta+\eta} = \{\mathbf{x}(t_{u_i}) : u_i(\zeta^\eta) > 0 \wedge |t_\zeta + \eta - t_{u_i}| > w\}_{i=1}^{N_n}, \quad [5]$$

in analogy to what was done for the reference state ζ in Eq. 3.

With the three sets just introduced in Eqs. 3–5, namely the recurrences R_{t_ζ} in Eq. 3, the forward recurrences $R_{t_\zeta}^\eta$ in Eq. 4, and the forward-reference-state recurrences $R_{t_\zeta+\eta}$ in Eq. 5, we can define the local predictability index as follows:

$$\alpha_\eta(\zeta) = \frac{|R_{t_\zeta}^\eta \cap R_{t_\zeta+\eta}|}{|R_{t_\zeta}|}, \quad [6]$$

where $|\cdot|$ represents the number of elements within the set (\cdot) , and \cap is the intersection of two sets. Since both R_{t_ζ} and $R_{t_\zeta+\eta}$ are defined using the same quantile-based thresholds, the number of elements in one set is equal to the number of elements in the other set, that is, $|R_{t_\zeta}| = |R_{t_\zeta+\eta}| = N_n$. In addition, given that $R_{t_\zeta}^\eta$ is defined from R_{t_ζ} , the number of elements it contains is also equal to N_n . It follows that $0 \leq \alpha_\eta(\zeta) \leq 1$. A value of $\alpha_\eta(\zeta)$ close to 1 indicates high probability for a neighbor state to remain close to the reference state. We interpret this as a high predictability, in a statistical sense, for state ζ at time t_ζ , and forecasting horizon η , suggesting that most of its neighboring states stay close to the future trajectory of ζ as time evolves. Conversely, when $\alpha_\eta(\zeta)$ is close to zero, this implies that most of the initial neighbors have moved away from the trajectory of ζ after time η , indicating low predictability in a statistical sense.

The derivation of the new quantity $\alpha_\eta(\zeta)$ relies on the distance function dist and the quantile q , both of which are used to define recurrences. Although the Euclidean distance in phase space has been shown to be an appropriate choice for physical systems (47), the distance function can be adjusted depending on the prior knowledge of the system and the application context. The selection of the quantile q can be adapted for different application scenarios, and it reflects the local statistical predictability for different scales. This is not surprising as we are defining α_η using extreme value theory where quantile choice is always an important step. In *SI Appendix, section 2*, we show an example of this scale-dependent behavior using various q . This dependence on q can be seen as a favorable property as it allows for analyzing predictability at different scales.

In addition, the new quantity $\alpha_\eta(\zeta)$ is inherently connected to information theory, since estimating local predictability can be viewed as a process that quantifies the uncertainty or randomness within the system. In fact, if we take a conditional probability perspective, where $P(A|B)$ represents the conditional probability of A given B , then we can reformulate the definition of α_η in Eq. 6 as follows:

$$\alpha_\eta(\zeta) = P[\mathbf{x}(t + \eta) \in R_{t_\zeta+\eta} | \mathbf{x}(t) \in R_{t_\zeta}], \quad [7]$$

where $A = \mathbf{x}(t + \eta) \in R_{t_\zeta+\eta}$ corresponds to recurrences of the reference state ζ that, after being propagated forward in time by η , belong to the forward-state neighborhood ζ^η (i.e., they are forward-reference-state recurrences), and $B = \mathbf{x}(t) \in R_{t_\zeta}$ corresponds to the recurrences of the reference state ζ . We can further link α_η to information theory using the Shannon entropy, when defined—see *SI Appendix, section 3*.

The new predictability index has a number of advantages compared to existing predictability measures: i) it is purely data-driven, ii) it is local as it measures instantaneous properties, iii) it

does not require simplifications to be made on the data, and iv) it is suitable for high dimensional datasets.

However, given that it uses future states, it cannot be readily used for real-time predictability analyses, for instance in the context of operational weather forecasting. We shall discuss this aspect in a later section and propose a practical way to address this limitation.

It is helpful here to emphasize some key differences to other data-driven ways of measuring local predictability properties. The classical algorithms proposed in ref. 53 to measure the Lyapunov exponents from available time series allow to estimate the growth rate of the n -dimensional volumes (for n smaller than or equal to the dimension of the phase space) defined by $n + 1$ initially nearby points. The strategy there strongly relies on using delay coordinates in order to take advantage of Takens' theorem (54) and reconstruct the key properties of the attractor. By selecting nearby points belonging to a region of interest, local properties can be extracted. Following such geometrical construction, one obtains that

$$P[\text{dist}(\mathbf{x}(t + \eta), \zeta^\eta) < \epsilon | \text{dist}(\mathbf{x}(t), \zeta) < \epsilon] \approx e^{-\left(\sum_{\lambda_j > 0} \lambda_j^\zeta\right)\eta}, \quad [8]$$

where the exponent is proportional to the sum of positive local, finite-time, Lyapunov exponents. Clearly, ϵ is assumed to be small. After comparing Eqs. 7 and 8 it is tempting to deduce

$\alpha_\eta \approx e^{-\left(\sum_{\lambda_j > 0} \lambda_j^\zeta\right)\eta}$ (see later discussion on specific numerical examples). Yet, a key difference between the two definitions exists when one applies them to actual data, because in Eq. 7 one considers neighborhoods of ζ and ζ^η having the same number of points (thanks to the use of quantiles in defining the threshold of the distance observables), so that such neighborhoods will in general have different radii (which are instead identical and equal to ϵ in Eq. 8).

Results and Discussion

We have computed the new metric $\alpha_\eta(\zeta)$ for systems of varying complexity to show its effectiveness. The following subsections provide a detailed discussion of the results for the Lorenz-63 attractor and climate data characterizing atmospheric circulation in the Euro-Atlantic sector. Additionally, a brief summary of $\alpha_\eta(\zeta)$ applied to other systems is presented in a subsequent section to illustrate its broad applicability. Following the discussion of results obtained from various systems, we propose a method for utilizing this index in real-time, addressing the issue introduced at the end of the previous section.

Lorenz-63. We consider the Lorenz-63 attractor as an example of idealized dynamical system. The data for this test case is obtained from the numerical simulation of the Lorenz-63 system, with a time resolution of 5×10^{-3} over 1×10^5 time steps, using the classical set of parameters that leads the system to be chaotic (see *SI Appendix, section 1* for details). Fig. 2 shows the distribution of α_η for the Lorenz-63 system, for five different forecasting horizons: $\eta = [0.05, 0.1, 0.2, 1, 2]\eta_\ell$. The quantity η_ℓ , that represents the Lyapunov time, is equal to 1.1 time units and is computed as the inverse of the maximum Lyapunov exponent of the Lorenz-63 system.

Fig. 2 A–E depict the Lorenz attractor in the phase space colored by the values of α_η , for the 5 different forecasting horizons considered. For each panel in Fig. 2 A–E, the title on

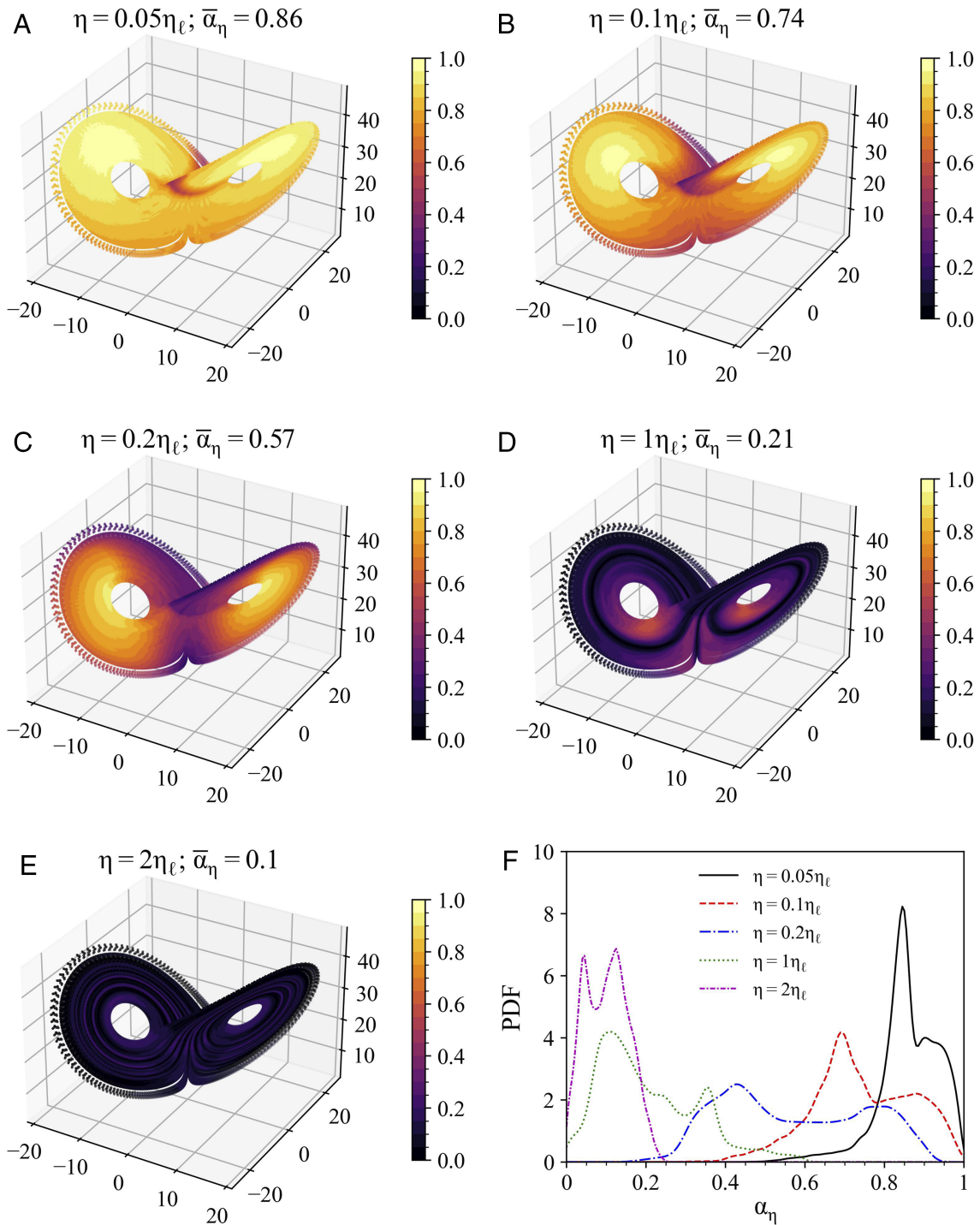


Fig. 2. Distribution of α_η at different prediction horizons for the Lorenz-63 system. (A–E) Lorenz attractor colored by α_η values at five different forecasting horizons: $\eta = [0.05, 0.1, 0.2, 1, 2]\eta_\ell$, each corresponding to timesteps: $L = [11, 22, 44, 220, 440]$. (F) Probability distribution of α_η at the same five different forecasting horizons. The quantile q applied in this analysis is 0.99, and the Theiler window size w is set to 50 time steps.

the top reports the forecasting horizon, and the average of the local predictability, namely $\bar{\alpha}_\eta$. We note how, as we increase the forecasting horizon, $\bar{\alpha}_\eta$ tends to decrease, as also depicted by the increasingly darker colors from Fig. 2 A–E. This result is in agreement with what we shall expect—the longer the forecasting horizon, the less predictable the system. This conclusion is also summarized in Fig. 2F, that shows the distribution of α_η for the five different forecasting horizons: The two longest horizons display a distribution peaked on low values of predictability (~ 0.2), while shorter horizons tend to display distributions

peaked toward higher values of predictability. To further test the resilience of α_η to noise, we applied α_η to the Lorenz-63 system with stochastic diffusion terms for comparison (SI Appendix, Fig. S1). The results are similar to Fig. 2, while we observe small-scale fluctuations of α_η over the attractor and a more rapid drop in average predictability, as one shall expect.

These results confirm that the predictability of the Lorenz-63 system is closely related to the local reference state in the phase space and the forecasting horizon. In addition, the phase-space distribution of α_η shows some interesting features,

A Area Under Curve

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|------------------------|-------|-------|-------|-------|-------|-------|-------|
| $\eta = 0.05\eta_\ell$ | 0.047 | 0.049 | 0.045 | 0.045 | 0.045 | 0.049 | 0.047 |
| $\eta = 0.1\eta_\ell$ | 0.094 | 0.099 | 0.086 | 0.085 | 0.086 | 0.099 | 0.094 |
| $\eta = 0.2\eta_\ell$ | 0.184 | 0.184 | 0.15 | 0.139 | 0.151 | 0.186 | 0.185 |
| $\eta = 1\eta_\ell$ | 0.363 | 0.294 | 0.328 | 0.229 | 0.336 | 0.309 | 0.374 |
| $\eta = 2\eta_\ell$ | 0.473 | 0.394 | 0.515 | 0.35 | 0.544 | 0.44 | 0.522 |
| $\eta = 5$ | 0.596 | 0.502 | 0.728 | 0.485 | 0.787 | 0.573 | 0.681 |

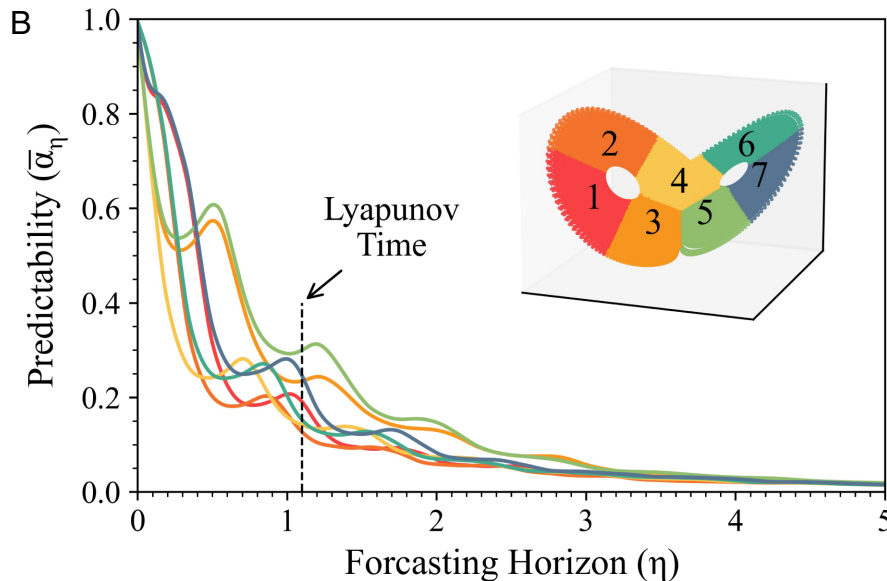


Fig. 3. Temporal and phase-spatial variations of α_η . (A) Area under the curve for different forecasting horizons and phase space regions, calculated from the time series shown in panel (B). Each cell color corresponds to the respective cluster in (B). The evolution of α_η averaged across seven clusters of states on the Lorenz-63 attractor. Inset: Lorenz-63 attractor colored to represent seven clusters obtained from K-means algorithm, each labeled with corresponding numbers.

namely: for short forecasting horizons the region where the Lorenz-63 attractor switches lobes seems to have low predictability (Fig. 2A), while for longer forecasting horizons, also the wings of the attractor become less predictable (Fig. 2B and C). These findings are consistent with previous studies on local predictability of Lorenz systems (21, 27). Remarkably, for forecasting horizons twice the Lyapunov time, the overall system loses predictability (Fig. 2E).

To further show how α_η reveals phase-space and time features that can be used for understanding the predictability of dynamical systems, we present the evolution of the average α_η for seven different clusters (or regions) of the Lorenz-63 attractor. These clusters are depicted in Fig. 3, and are obtained by applying a K-means clustering method to the Lorenz-63 data used for Fig. 2, similar to the one adopted in ref. 55. Interestingly, the clustering identifies portions of the attractor in the phase space that seem to match some of the observations made for Fig. 2. In particular, cluster 4 represents the region where trajectories are switching lobes, clusters 1, 2, 6, and 7 represent the wings of the attractor, while clusters 3 and 5 represent intermediate regions that contain features from the five clusters just described. We note how, as expected, cluster 4 has the lowest short-term predictability among the regions considered, with a sudden drop at short forecasting horizons (yellow line). The wings of the attractor follow, with clusters 2 (dark orange) and 6 (green)

being the least predictable after cluster 4, immediately followed by clusters 1 (red) and 7 (blue). Clusters 3 and 5, after a sudden predictability drop similar to the one experienced for cluster 4, maintain relatively good predictability for larger forecasting horizons compared to all other clusters. This is confirmed, if we look at the area under each curve (AUC) reported in Fig. 3, that represent the average predictability of each cluster with respect to the forecasting horizons, with clusters 3 and 5 having the highest average predictability, and cluster 4 the lowest.

We also note that, despite the general trend of predictability decreasing as the prediction horizon increases, α_η does not always decrease monotonically. There are instances showing returns of predictability, as also reported in ref. 56, and as reflected more clearly in Fig. 3B; see also discussion above regarding the relationship between the definitions provided in Eqs. 7 and 8.

We additionally compare our results with the nonlinear local Lyapunov exponent (NLLE) (36, 37), an established index used to quantify local predictability that we find closely relevant, as reported in *SI Appendix, section 4*. The results for the NLLE display some notable differences with our local predictability index while also showing some similarities (*SI Appendix, Fig. S6*). These differences are arguably due to the use of different perspectives in defining predictability. The NLLE framework focuses on the exponential error growth rate, while our method is rooted in the concept of recurrences in dynamical system

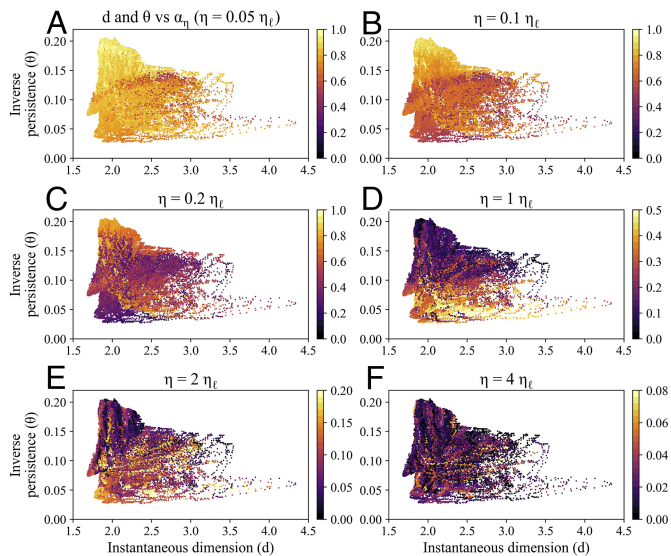


Fig. 4. Comparative analysis of α_η with local dynamical indices for the Lorenz-63 system. (A–F) Scatter plots of local dimension versus local inverse persistence, where each dot represents one state. All states are colored based on α_η values for $\eta = [0.05, 0.1, 0.2, 1, 2, 4]\eta_\ell$. Note that the color scale is different in the last three panels in order to improve readability.

theory. We acknowledge that both methods have their own caveats: NLE predefines the rule of error growth, while our method neglects the information on trajectories that have left the neighborhood. To overcome this, we propose an extension of our predictability index by accounting for the information given by the distance of all the neighboring trajectories—i.e., both those that remained close to the reference state and those that departed from it, as shown in *SI Appendix, section 5*.

Furthermore, we evaluated the relationship of the previously introduced local dynamical indices, d and Θ , to predictability, by comparing them against α_η . In Fig. 4, we show scatter plots of the local predictability index α_η versus the local dimension d and the inverse persistence θ . We note that the intuition that high dimension and low persistence means low predictability does not hold for the case of the Lorenz-63 system. Specifically, states on the two wings of the Lorenz attractor exhibit low persistence (high values of θ), indicating that these states are likely to leave the neighborhood immediately (50). However, they also demonstrate high predictability over short forecasting horizons. This is not surprising when considering the dynamics on the wings of the Lorenz attractor—although these states may leave the neighborhood quickly, their neighboring trajectories do not diverge significantly, making them predictable in short forecasting horizons.

We show, in the next section, that, instead, for the real-world data considered, the agreement between the interpretation of the d - θ space and α_η seems to hold.

Atmospheric Circulation. In addition to the test case presented for idealized systems, we computed α_η using real-world data. In particular, we focused on studying the predictability of large-scale atmospheric circulation patterns in the Euro-Atlantic sector (see *Materials and Methods* for details). In this context, we consider the Z500 map for each day as a state along the system’s trajectory in the approximated phase space. We emphasize that this region serves as an ideal testbed to validate our predictability index due to its extensive body of relevant research, especially considering that no state-dependent predictability index is directly comparable for high-dimensional datasets.

In the Euro-Atlantic sector, the definition of weather regimes offers a direct way of classifying atmospheric states.

Weather regimes are defined as recurrent and quasi-stationary large-scale circulation patterns (57, 58). Despite being a coarse characterization of atmospheric variability, weather regimes are found to be associated with conditions that are specifically favorable for certain types of weather extremes (59), and they have exhibited varying levels of predictability in real-time forecasting (60). In this work, we have adopted the four-regime definition widely used for scientific research (61), although various definitions for weather regimes exist (61, 62). Based on this definition, all extended wintertime (DJFM) Z500 daily patterns are classified into five different categories including NAO+, Atlantic Ridge (AR), NAO–, Scandinavian Blocking (SB), and no regime, using a method slightly adapted from ref. 63 (see *SI Appendix, section 6* for details).

Fig. 5 shows the distribution of α_η across all five categories for $\eta = 1$ to 3 d. The results indicate that the NAO– regime has the highest predictability, followed by NAO+, with the remaining two regimes showing no significant difference from having no regime for $\eta = 1$ d. For $\eta = 2$ and 3 d, NAO+ and NAO– still maintain notably higher mean values compared to other categories, with the no regime category having the lowest values, confirming that large scale weather regimes are a primary source of predictability in the Euro-Atlantic region. These results are in excellent agreement with previous investigation dedicated to the Euro-Atlantic sector: NAO– exhibits stronger predictability than other regimes, while the blocked zonal regimes (SB and AR) are notoriously unpredictable (50, 60, 64). Additionally, we find further confirmation of the general fact that blockings are associated with anomalously high instability compared to zonal flow conditions (65, 66). These findings further validate the effectiveness of our index, supported by both the underlying physical explanations and the model-based prediction results. We note that the analysis just presented can be extended to any other variables and domains, with a key region of interest being the

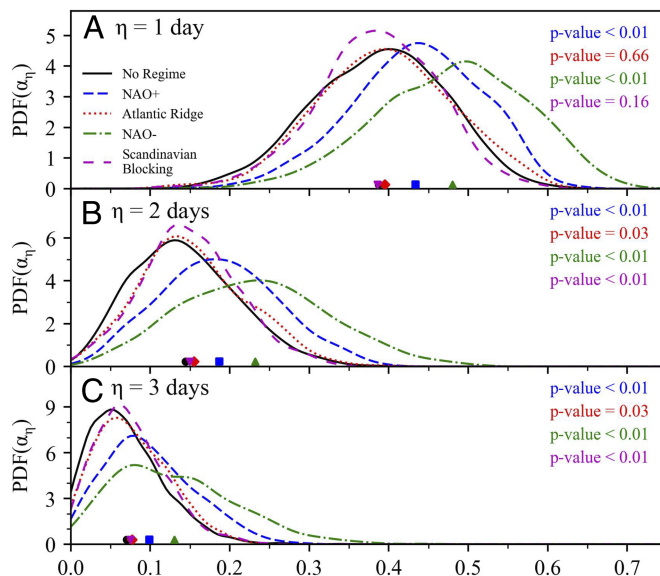


Fig. 5. Predictability analysis of North Atlantic wintertime weather regimes. (A–C) Probability distribution of α_η for four weather regime and no regime: NAO+ (blue dashed line), Atlantic Ridge (red dotted line), NAO– (green dot-dashed line), Scandinavian Blocking (magenta loosely dashed line), and no regime (black line), for $\eta = 1$ to 3 d. The markers on the x-axis show the mean value of α_η for different categories. The statistical significance is evaluated between the distribution of each weather regime and that of no regime using the Kolmogorov-Smirnov test. The quantile q applied in this analysis is 0.99, and the Theiler window size w is set to 7 d.

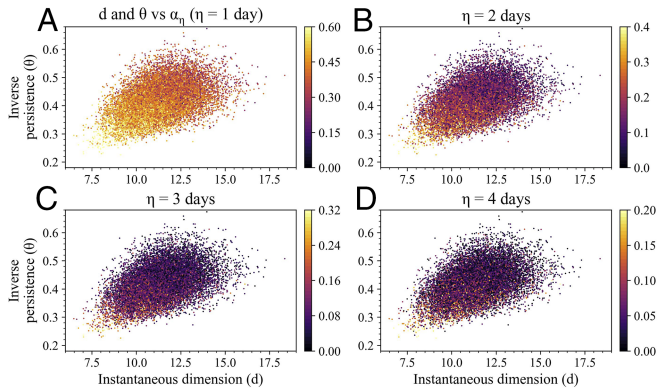


Fig. 6. Comparative analysis of α_η with local dynamical indices for Z500 patterns in the Euro-Atlantic sector. (A–D) Scatter plots of local dimension versus local inverse persistence, where each dot represents one daily Z500 pattern. All states are colored based on α_η values for $\eta = 1$ to 4 d. Note that the color scale is different in each panel in order to improve readability.

tropical Indo-Pacific, given the complex interaction of different modes of variability (67).

Fig. 6 illustrates scatter plots of local dimension d vs. inverse persistence θ for all the atmospheric states, colored by the predictability index α_η , for different prediction horizons (η) ranging from 1 to 4 d. The overall predictability drops as η increases, as expected due to the highly chaotic nature of atmospheric circulation. Notably, states in the lower-left corner of the d - θ scatter plot space consistently exhibit higher-than-average predictability, regardless of the value of η . This aligns with previous results suggesting that states with low local dimension and high persistence are more predictable. From a dynamical system theory perspective, a lower local dimension indicates a simpler dynamics, while higher persistence means that a state is likely to remain similar to its current configuration for a longer time. This result demonstrates the effectiveness of the proposed predictability index for high-dimensional systems.

Other Systems. In addition to the Lorenz-63 system and climate data, in Fig. 7, we demonstrate the applicability of α_η across a diverse range of systems. These include another canonical dynamical system (Rössler attractor), simulation data for slow earthquakes (spring-slider system), a classical physics problem (double pendulum), and real-world biological data (observational ECG and EEG data). The stochastic spring-slider system was originally introduced to simulate slow earthquakes, which feature a random attractor governed by stochastic differential equations, as detailed in ref. 16. The final three examples use the same dataset as that of ref. 68, with the double pendulum data generated through numerical simulation, while the other two are obtained from observational data.

The results for the Rössler attractor show a certain degree of similarity to those of the Lorenz-63 system: The Lorenz-63 system experiences a significant drop in predictability before lobe transitions, whereas the Rössler attractor exhibits a similar phenomenon prior to bursting events in the z -direction. We also note that the region following the bursting events exhibits markedly higher predictability, as the trajectories passing through this region will oscillate predictably in the x - y plane until reentering the region prior to the bursting events. In the case of the spring-slider system, there is high predictability during the almost linear interseismic period. The time-lagged recurrence index starts decreasing when the slip rate of the slider increases.

The fact that the onset of the rupture is marked by high unpredictability is in line with the idea that it is a difficult task to predict the final size of the event at the onset of the slipping phase. For the attractor of the double pendulum, we note that two endpoints of the attractor show noticeably lower predictability, whereas the transitional regions between them show relatively higher predictability, particularly for the outer parts of the attractor. In the two real-world biological data examples, we also observe that α_η reveals variability in predictability within the phase space across different forecasting horizons. These insights can improve the analysis of such systems and deepen our understanding of their underlying dynamics. The applications to these diverse systems collectively demonstrate that this method has broad applicability and effectiveness.

A Note on Real-Time Predictability. Our findings demonstrate the agreement of α_η with known results, positioning it as a viable index for analyzing the local predictability of dynamical systems. Yet, one key point remains outstanding: α_η , as is, cannot be used for real-time applications, as it requires the future state of the system. However, since α_η quantifies local predictability information in the phase space, for a given unprecedented state ζ' , we can compute the average α_η of its analogues (i.e., similar states to ζ' that appeared in the past). The number of analogues can be set once again using a quantile threshold, but this could be different from the one used to define the local properties. In practice, if we consider N_a analogues, indicated by $\{\mathbf{x}(t_{a_i})\}_{i=1}^{N_a}$, with t_{a_i} being the time at which the analogue number a_i occurred,

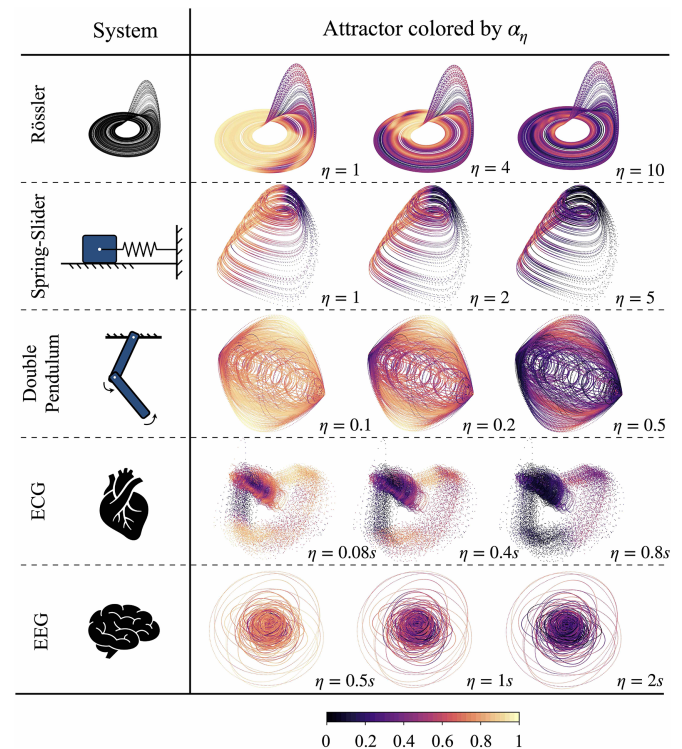


Fig. 7. Summary of α_η applied to various systems, including another canonical dynamical system (the Rössler attractor), simulated slow earthquakes (spring-slider model), a classical physics problem (the double pendulum), and real-world biological data (observational ECG and EEG data). The forecasting horizon η is annotated at the Bottom-Right corner of each attractor. All the attractors presented here are three dimensional for demonstration purposes, but their corresponding α_τ are computed in higher-dimensional spaces, as detailed in *SI Appendix, section 7*. The quantile q applied in all analysis is 0.99, and the Theiler window size w is set to 50 time steps.

the predictability index proxy is

$$\hat{\alpha}_\eta(\zeta') = \frac{1}{N_a} \sum_{i=1}^{N_a} \alpha_\eta(\mathbf{x}(t_{a_i})) \quad [9]$$

This can be used for real-time predictability analyses of dynamical systems. For instance, in the context of weather and climate applications, we can identify the analogues of today's data, and compute the predictability index proxy $\hat{\alpha}_\eta(\zeta')$ using Eq. 9. This will give us the real-time predictability of the system as of today, thereby allowing us to adopt the index for real-time applications. An example of real-time applications is presented in [SI Appendix, section 8](#), where we show an application for the Lorenz-63 system, with the real-time predictability proxy $\hat{\alpha}_\eta(\zeta')$ successfully capturing the phase-space and time features of the predictability.

Conclusions

This work complements the rich body of work on the pivotal topic of predictability of dynamical systems, from the pioneering works of Poincaré and Lorenz, to modern dynamical system theory and its applications to real-life complex systems, especially in the context of weather and climate prediction. In particular, we provide a time-lagged recurrence index that can infer state-dependent predictability from high-dimensional observational datasets, overcoming the curse of dimensionality that affects more traditional approaches. While our approach has resemblances with classical data-driven methods used for performing Lyapunov analysis on general systems (53, 69), the local predictability index differs from these methods and is able to describe flexibly key dynamical features of complex systems. We applied this index to dynamical systems of varying complexity, ranging from idealized low-dimensional systems to real-world high-dimensional atmospheric reanalysis datasets and we showcased its effectiveness. In particular, our findings confirm that when in the Euro-Atlantic sector of the mid-latitudes a blocking is present, the predictability of the atmosphere is typically anomalously low. We further showed how the new index reveals the scale-dependent nature of predictability by using different quantiles, and how it can be used in real-time problems. Our approach also addresses the ambiguity in terms of predictability that surrounds the two local dynamical indices, namely dimension d and persistence Θ , providing a framework that directly addresses predictability.

We conclude this work with a note. The definition of α_η does not require $\eta > 0$. Therefore, α_η can be computed backward in time using $\eta < 0$. In this case, α_η characterizes the dynamics of the system as it approaches the reference state of interest, and it provides information that is entirely different from the forward-in-time analysis presented so far, using $\eta > 0$. Indeed, the backward predictability analysis can be used for understanding the dynamical pathways leading to extreme events as well as other phenomena in dynamical systems (70, 71). As a methodological study, we consider such in-depth analysis to be beyond the scope of this work, but we believe it may shed new light on various scientific domains.

Materials and Methods

Local Dynamical Indices. In the framework proposed by Lucarini et al. (47), two dynamical indices, the local dimension d and persistence Θ , were introduced to characterize the local properties of dynamical systems. To compute these indices, we first consider the negative logarithmic returns $g(\zeta, \mathbf{x}(t))$, as defined in Eq. 1, which quantifies proximity to a given state ζ . The exceedance $\mathbf{u}(\zeta)$,

introduced in Eq. 2, is defined for the neighboring states (i.e., recurrences) of ζ and measures how much $g(\zeta, \mathbf{x}(t))$ exceeds a high quantile q . According to extreme value theory (72, 73), assuming the independence of the exceedances, the cumulative probability distribution $F(\mathbf{u}, \zeta)$ follows the exponential member of the generalized Pareto distribution (GPD), that is

$$F(\mathbf{u}, \zeta) \sim \exp \left[-\frac{\mathbf{u}(\zeta)}{\sigma(\zeta)} \right]. \quad [10]$$

The parameter $\sigma(\zeta)$, which is the scale parameter of the distribution, depends on the state ζ in phase space and can be used to derive the local dimension via $d(\zeta) = 1/\sigma(\zeta)$ (47). In practice, the local dimension is computed by explicitly fitting an exponential distribution to the exceedances $\mathbf{u}(\zeta)$; see discussion in refs. 74 and 75. The estimate of the attractor's dimension derived from $d(\zeta)$ seems to be more efficient than previous methods based on evaluating correlation integrals (76, 77). The other local index, persistence (Θ), is defined as the inverse of the extremal index $0 \leq \theta \leq 1$ (78, 79), a dimensionless parameter that measures the inverse of the length of clustering of extremes. In this context, extremes are defined as the *recurrences* within the neighborhood around a given state in phase space. Therefore, persistence (Θ) estimates the average number of consecutive recurrences, namely the time steps for the system to resemble a given state in its neighboring phase space.

This framework is the basis for the predictability index proposed in this work, and introduced in the main text above.

500 hPa Geopotential Data in the Euro-Atlantic Sector. In this study, the atmospheric circulation dynamics in the Euro-Atlantic sector (80°W to 50°E, 22.5°N to 70°N) served as a testbed to evaluate the applicability of α_η on a real-world high-dimensional dataset. We used the state-of-the-art ERA5 reanalysis data (80) from 1979 to 2022, with a spatial resolution of 0.25°. The daily mean 500 hPa geopotential height (Z500) was selected as the representative variable, which has been frequently used to characterize atmospheric circulation dynamics in the mid-latitudes (58, 81–84). For this spatiotemporal system, we treat each daily spatial map of Z500 as an individual state, and the corresponding space can be regarded as an approximation to the phase space. This approach is widely adopted in practical applications involving real-world complex systems because accessing all variables of a system to construct its full phase space is often impossible. Instead, we can focus on a subset of available observables to approximate the system's dynamics. In this way, the evolution of atmospheric circulation patterns can be viewed as a trajectory within this high-dimensional space, thereby enabling the application of α_η to assess its state-dependent predictability.

Data, Materials, and Software Availability. The code for computing the local predictability index and the data used in this study are available at <https://github.com/MathEXLab/TLR> (85). The ERA5 reanalysis data used in this study are available at <https://cds.climate.copernicus.eu/#/search?text=ERA5&type=dataset> (80) The data used for the double pendulum, ECG, and EEG are from ref. 68.

ACKNOWLEDGMENTS. G.M. acknowledges support from Singapore's Ministry of Education Tier 2 grant no T2EP50221-0017, titled: "Prediction-to-Mitigation with Digital Twins of the Earth's Weather." V.L. acknowledges the support provided by the Horizon Europe Project ClimTIP (Grant No. 100018693) and by the UK's Engineering and Physical Sciences Research Council project LINK (Grant No. EP/Y026675/1).

Author affiliations: ^aDepartment of Mechanical Engineering, National University of Singapore, Singapore 117575, Singapore; ^bLaboratoire des Sciences du Climat et de l'Environnement, Commissariat à l'énergie atomique et aux énergies alternatives Saclay l'Orme des Merisiers Unité Mixte de Recherche 8212 CEA-Centre National de la Recherche Scientifique-Université de Versailles - Saint-Quentin-en-Yvelines, Université Paris-Saclay & Institut Pierre Simon Laplace, 91191, Gif-sur-Yvette, France; ^cLondon Mathematical Laboratory, London W6 8RH, United Kingdom; ^dLaboratoire de Météorologie Dynamique/IPSL, École Normale Supérieure, Paris Sciences et Lettres Research University, Sorbonne Université, École Polytechnique, Paris, CNRS, Paris 75005, France; ^eBullard Laboratories, Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, United Kingdom; ^fOsservatorio Nazionale Terremoti, Istituto Nazionale di Geofisica e Vulcanologia, Rome 00143, Italy; ^gSchool of Computing and Mathematical Sciences, University of Leicester, Leicester, LE1 7RH United Kingdom; and ^hDepartment of Mathematics, National University of Singapore, Singapore 119076, Singapore

1. I. Newton, *Philosophiæ Naturalis Principia Mathematica* (G. Brookman, 1833), vol. 1.
2. H. Poincaré, Sur le problème des trois corps et les équations de la dynamique. *Acta Math.* **13**, A3–A270 (1890).
3. S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, 2018).
4. J. P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617 (1985).
5. A. Pikovsky, A. Politi, *Lyapunov Exponents: A Tool to Explore Complex Dynamics* (Cambridge University Press, 2016).
6. H. Risken, T. Frank, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer Series in Synergetics (Springer, Berlin Heidelberg, 2012).
7. G. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations, Texts in Applied Mathematics* (Springer, New York, 2014).
8. K. Hasselmann, Stochastic climate models Part I. Theory. *Tellus* **28**, 473–485 (1976).
9. V. Lucarini, M. D. Chekroun, Theoretical tools for understanding the climate crisis from Hasselmann's programme and beyond. *Nat. Rev. Phys.* **5**, 744–765 (2023).
10. H. Mori, Transport, collective motion, and Brownian motion. *Progr. Theoret. Phys.* **33**, 423–455 (1965).
11. R. Zwanzig, Memory effects in irreversible thermodynamics. *Phys. Rev.* **124**, 983–992 (1961).
12. M. Santos Gutiérrez, V. Lucarini, M. D. Chekroun, M. Ghil, Reduced-order models for coupled dynamical systems: Data-driven methods and the Koopman operator. *Chaos* **31**, 053116 (2021).
13. C. Penland, T. Magorian, Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Clim.* **6**, 1067–1076 (1993).
14. D. Kondrashov, S. Kravtsov, M. Ghil, Empirical mode reduction in a model of extratropical low-frequency variability. *J. Atmos. Sci.* **63**, 1859–1877 (2006).
15. D. Faranda *et al.*, Stochastic chaos in a turbulent swirling flow. *Phys. Rev. Lett.* **119**, 014502 (2017).
16. A. Gualandri, D. Faranda, C. Marone, M. Cocco, G. Mengaldo, Deterministic and stochastic chaos characterize laboratory earthquakes. *Earth Planet. Sci. Lett.* **604**, 117995 (2023).
17. S. Nath, A. Raj, K. Thingbaijam, A. Kumar, Ground motion synthesis and seismic scenario in Guwahati city—a stochastic approach. *Seismol. Res. Lett.* **80**, 233–242 (2009).
18. L. T. Giorgini, W. Moon, N. Chen, J. S. Wettlaufer, Non-gaussian stochastic dynamical model for the El Niño southern oscillation. *Phys. Rev. Res.* **4**, L022065 (2022).
19. E. N. Lorenz, Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).
20. S. Siebert, H. Kantz, Prediction of complex dynamics: Who cares about chaos? *Chaos Detect. Predict.* **915**, 249–269 (2016).
21. T. N. Palmer, Extended-range atmospheric prediction and the Lorenz model. *Bull. Am. Meteorol. Soc.* **74**, 49–66 (1993).
22. H. Crauel, A. Debussche, F. Flandoli, Random attractors. *J. Dyn. Diff. Equat.* **9**, 307–341 (1997).
23. G. Boffetta, M. Cencini, M. Falconi, A. Vulpiani, Predictability: A way to characterize complexity. *Phys. Rep.* **356**, 367–474 (2002).
24. U. Parlitz, Estimating Lyapunov exponents from time series. *Chaos Detect. Predict.* **16**, 1–34 (2016).
25. V. I. Oseledec, A multiplicative ergodic theorem, Lyapunov characteristic numbers for dynamical systems. *Trans. Mosc. Math. Soc.* **19**, 197–231 (1968).
26. C. C. Maiocchi, V. Lucarini, A. Gritsun, Y. Sato, Heterogeneity of the attractor of the Lorenz '96 model: Lyapunov analysis, unstable periodic orbits, and shadowing properties. *Phys. D Nonlinear Phenom.* **457**, 133970 (2024).
27. J. M. Nese, Quantifying local predictability in phase space. *Phys. D Nonlinear Phenom.* **35**, 237–250 (1989).
28. C. Ziehmann, L. A. Smith, J. Kurths, Localized Lyapunov exponents and the prediction of predictability. *Phys. Lett. A* **271**, 237–251 (2000).
29. J. Slingo, T. Palmer, Uncertainty in weather and climate prediction. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **369**, 4751–4767 (2011).
30. A. Carrassi, M. Bocquet, L. Bertino, G. Evensen, Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev. Clim. Chang* **9**, e535 (2018).
31. H. D. Abarbanel, R. Brown, M. B. Kennel, Variation of Lyapunov exponents on a strange attractor. *J. Nonlinear Sci.* **1**, 175–199 (1991).
32. B. Eckhardt, D. Yao, Local Lyapunov exponents in chaotic systems. *Phys. D Nonlinear Phenom.* **65**, 100–108 (1993).
33. M. Cencini, A. Vulpiani, Finite size Lyapunov exponent: Review on applications. *J. Phys. A Math. Theor.* **46**, 254019 (2013).
34. E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, A. Vulpiani, Predictability in the large: An extension of the concept of Lyapunov exponent. *J. Phys. A Math. Gen.* **30**, 1 (1997).
35. G. Boffetta, P. Giuliani, G. Paladin, A. Vulpiani, An extension of the Lyapunov analysis for the predictability problem. *J. Atmos. Sci.* **55**, 3409–3416 (1998).
36. R. Ding, J. Li, Nonlinear finite-time Lyapunov exponent and predictability. *Phys. Lett. A* **364**, 396–400 (2007).
37. X. W. Huai, J. P. Li, R. Q. Ding, J. Feng, D. Q. Liu, Quantifying local predictability of the Lorenz system using the nonlinear local Lyapunov exponent. *Atmos. Oceanic Sci. Lett.* **10**, 372–378 (2017).
38. K. Geist, U. Parlitz, W. Lauterborn, Comparison of different methods for computing Lyapunov exponents. *Progr. Theoret. Phys.* **83**, 875–893 (1990).
39. T. DelSole, M. K. Tippett, Predictability: Recent insights from information theory. *Rev. Geophys.* **45**, RG000202 (2007).
40. T. Schneider, S. M. Griffies, A conceptual framework for predictability studies. *J. Clim.* **12**, 3133–3155 (1999).
41. R. Kleeman, Statistical predictability in the atmosphere and other dynamical systems. *Phys. D Nonlinear Phenom.* **230**, 65–71 (2007).
42. L. Y. Leung, G. R. North, Information theory and climate prediction. *J. Clim.* **3**, 5–14 (1990).
43. R. Kleeman, Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.* **59**, 2057–2072 (2002).
44. X. Li, G. Ouyang, D. A. Richards, Predictability analysis of absence seizures with permutation entropy. *Epilepsy Res.* **77**, 70–74 (2007).
45. O. J. Ikuyajolu, F. Falasca, A. Bracco, Information entropy as quantifier of potential predictability in the tropical Indo-pacific basin. *Front. Clim.* **3**, 675840 (2021).
46. C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **88**, 174102 (2002).
47. V. Lucarini *et al.*, *Extremes and Recurrence in Dynamical Systems* (John Wiley & Sons, 2016).
48. R. A. Fisher, L. H. C. Tippett, "Limiting forms of the frequency distribution of the largest or smallest member of a sample" in *Mathematical Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, 1928), vol. 24, pp. 180–190.
49. A. Gualandri, J. P. Avouac, S. Michel, D. Faranda, The predictable chaos of slow earthquakes. *Sci. Adv.* **6**, eaaz5548 (2020).
50. D. Faranda, G. Messori, P. You, Dynamical proxies of north Atlantic predictability and extremes. *Sci. Rep.* **7**, 41278 (2017).
51. A. Hochman, P. Alpert, T. Harpaz, H. Saaroni, G. Messori, A new dynamical systems perspective on atmospheric predictability: Eastern mediterranean weather regimes as a case study. *Sci. Adv.* **5**, eaau0936 (2019).
52. J. Theiler, Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A* **34**, 2427 (1986).
53. A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, Determining Lyapunov exponents from a time series. *Phys. D Nonlinear Phenom.* **16**, 285–317 (1985).
54. F. Takens, "Detecting strange attractors in turbulence" in *Dynamical Systems and Turbulence, Warwick 1980*, D. Rand, L. S. Young, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 1981), pp. 366–381.
55. E. Kaiser *et al.*, Cluster-based reduced-order modelling of a mixing layer. *J. Fluid Mech.* **754**, 365–414 (2014).
56. L. A. Smith, C. Ziehmann, K. Fraedrich, Uncertainty dynamics and predictability in chaotic systems. *Q. J. R. Meteorol. Soc.* **125**, 2855–2886 (1999).
57. P. A. Michelangeli, R. Vautard, B. Legras, Weather regimes: Recurrence and quasi stationarity. *J. Atmos. Sci.* **52**, 1237–1256 (1995).
58. S. Springer, A. Laio, V. M. Galfi, V. Lucarini, Unsupervised detection of large-scale weather patterns in the northern hemisphere via Markov state modelling: From blockings to teleconnections. *NPJ Clim. Atmos. Sci.* **7**, 105 (2024).
59. P. You, M. Nogaj, Extreme climatic events and weather regimes over the north Atlantic: When and where? *Geophys. Res. Lett.* **31**, GLO19119 (2004).
60. L. Ferranti, S. Corti, M. Janousek, Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Q. J. R. Meteorol. Soc.* **141**, 916–924 (2015).
61. R. Vautard, Multiple weather regimes over the north Atlantic: Analysis of precursors and successors. *Mon. Weather Rev.* **118**, 2056–2081 (1990).
62. C. M. Grams, R. Beerli, S. Pfenninger, I. Staffell, H. Wernli, Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nat. Clim. Change* **7**, 557–562 (2017).
63. C. Cassou, Intraseasonal interaction between the Madden-Julian oscillation and the north Atlantic oscillation. *Nature* **455**, 523–527 (2008).
64. R. Owens, T. Hewson, *ECMWF Forecast User Guide* (ECMWF, Reading, 2018).
65. S. Schubert, V. Lucarini, Dynamical analysis of blocking events: Spatial and temporal fluctuations of covariant Lyapunov vectors. *Q. J. R. Meteorol. Soc.* **142**, 2143–2158 (2016).
66. V. Lucarini, A. Gritsun, A new mathematical framework for atmospheric blocking events. *Clim. Dyn.* **54**, 575–598 (2020).
67. C. Dong *et al.*, Indo-pacific regional extremes aggravated by changes in tropical weather patterns. *Nat. Geosci.* **17**, 979–986 (2024).
68. S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, J. N. Kutz, Chaos as an intermittently forced linear system. *Nat. Commun.* **8**, 19 (2017).
69. M. T. Rosenstein, J. J. Collins, C. J. De Luca, A practical method for calculating largest Lyapunov exponents from small data sets. *Phys. D Nonlinear Phenom.* **65**, 117–134 (1993).
70. X. Li, R. Ding, The backward nonlinear local Lyapunov exponent and its application to quantifying the local predictability of extreme high-temperature events. *Clim. Dyn.* **60**, 2767–2781 (2023).
71. R. Noyelle, P. You, D. Faranda, Investigating the typicality of the dynamics leading to extreme temperatures in the IPSL-CM6A-LR model. *Clim. Dyn.* **62**, 1329–1357 (2024).
72. A. C. M. Freitas, J. M. Freitas, M. Todd, Hitting time statistics and extreme value theory. *Probab. Theory Relat. Fields* **147**, 675–710 (2010).
73. V. Lucarini, D. Faranda, J. Wouters, Universal behaviour of extreme value statistics for selected observables of dynamical systems. *J. Stat. Phys.* **147**, 63–73 (2012).
74. G. Datsiris, I. Kottlarz, A. P. Braun, U. Parlitz, Estimating fractal dimensions: A comparative review and open source implementations. *Chaos Interdiscip. J. Nonlinear Sci.* **33**, 102101 (2023).
75. F. Pons, G. Messori, D. Faranda, Statistical performance of local attractor dimension estimators in non-axiom a dynamical systems. *Chaos Interdiscip. J. Nonlinear Sci.* **33**, 073143 (2023).
76. P. Grassberger, I. Procaccia, Characterization of strange attractors. *Phys. Rev. Lett.* **50**, 346–349 (1983).
77. M. Ding, C. Grebogi, E. Ott, T. Sauer, J. A. Yorke, Estimating correlation dimension from a chaotic time series: When does plateau onset occur? *Phys. D Nonlinear Phenom.* **69**, 404–424 (1993).
78. A. C. M. Freitas, J. M. Freitas, M. Todd, The extremal index, hitting time statistics and periodicity. *Adv. Math.* **231**, 2626–2665 (2012).
79. N. R. Moloney, D. Faranda, Y. Sato, An overview of the extremal index. *Chaos Interdiscip. J. Nonlinear Sci.* **29**, 022101 (2019).
80. H. Hersbach *et al.*, The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
81. R. Benzi, P. Malguzzi, A. Speranza, A. Suter, The statistical properties of general atmospheric circulation: Observational evidence and a minimal theory of bimodality. *Q. J. R. Meteorol. Soc.* **112**, 661–674 (1986).
82. K. Mo, M. Ghil, Cluster analysis of multiple planetary flow regimes. *J. Geophys. Res. Atmos.* **93**, 10927–10952 (1988).
83. A. Dell'Aquila, V. Lucarini, P. M. Ruti, S. Calmanti, Hayashi spectra of the northern hemisphere mid-latitude atmospheric variability in the NCEP–NCAR and ECMWF reanalyses. *Clim. Dyn.* **25**, 639–652 (2005).
84. A. Jézéquel, P. You, S. Radanovics, Role of circulation in European heatwaves using flow analogues. *Clim. Dyn.* **50**, 1145–1159 (2018).
85. G. Mengaldo, TLR: Time-lagged recurrence (α_T). Github. <https://github.com/MathEXLab/TLR>. Deposited 10 February 2025.