

Dynamic Mode Decomposition of Extreme Events

Maša Ann¹, Jörn Behrens², and Jana Sillmann³

¹Universität Hamburg

²Universität Hamburg

³Universität Hamburg

Correspondence: Maša Ann (masa.ann@uni-hamburg.de)

Abstract. Most data-driven methods, among them Dynamic Mode Decomposition (DMD), focus on analysing and reconstructing the average ~~behavior~~behaviour of a system. However, the primary interest often lies in the anomalous behaviour, known as extreme events. This is especially the case in climate research, where extreme events have significant economic and societal costs. Therefore, we extend a DMD method to account for extreme events by adding a penalisation term. This extension allows us to not only better reconstruct the extreme events, but also extract the ~~spatio-temporal~~spatiotemporal structures related to those extreme events. DMD was originally developed by Schmid and Sesterhenn (Schmid and Sesterhenn, 2008) to enable the fluid dynamics community to identify ~~spatio-temporal~~spatiotemporal coherent structures (called *modes*) from high-dimensional data. In its essence DMD uses most relevant modes to filter the noise and reconstruct the original signal. We ask "Is the noise really noise"? Or can we attribute some of these dynamic modes, that result from the DMD, to extreme events? We applied this new method to the climate system, well known for its high-dimensionality. ~~We examined two heatwaves that occurred in Europe (HW-As a proof of concept, we applied the method to two well-studied European heatwaves: those of 2003 and HW-2010). In both cases we were able to improve the accuracy of the reconstruction. This novel variation of the DMD can also~~2010. Across both cases, our extreme DMD improves reconstruction accuracy at extreme spatiotemporal points, achieving a 0.45–0.85% relative reduction in error compared with standard DMD, a difference that is small in magnitude but statistically significant. The approach also reveals coherent spatial modes that contribute specifically to the development of heat extremes. This framework represents a general extension of DMD and can be applied to other ~~dynamical systems across many disciplines, in which~~high-dimensional dynamical systems where extreme events are of interest.

1 Introduction

Dynamic Mode Decomposition (DMD) is a relatively recent algorithm introduced by Schmid and Sesterhenn (Schmid and Sesterhenn, 2008). Originally developed for ~~modeling~~modelling fluids, DMD quickly developed as a powerful tool for ~~analyzing~~analysing the dynamics of nonlinear systems (H. Tu et al., 2014). It has also found its applications across many different disciplines (Brunton et al., 2021), e.g. neuroscience (Brunton et al., 2016), epidemiology (Proctor and Eckhoff, 2015), fluids mechanics (Schmid, 2010; Rowley et al., 2009b), video analysis (Erichson et al., 2016), climate dynamics (~~Froyland et al., 2021; Chowdary et al., 2023~~)(Froyland et al., 2021; Chowdary et al., 2023; Navarra et al., 2024; Zhang et al., 2025)

25 .

Recent studies have further expanded DMD's use in atmospheric science. For instance, Mankovich et al. (2025) applied the Dynamic Mode Decomposition with control (DMDc) framework to investigate how different anthropogenic emissions distinctly influence climate variability. DMD has also been applied to satellite observations, such as in Alcayaga et al. (2022), who identified large-scale atmospheric structures under varying stability conditions, and in Uchida et al. (2025), where DMD successfully separated spatial modes associated with sub-inertial and super-inertial periods.

The foundational concept of DMD is rooted in the well-established tradition of signal decomposition. DMD is a dimensionality reduction technique that transforms data from a high-dimensional space to a low-dimensional one while preserving the essential features of the original data. The resulting reconstruction typically captures the system's average behavior but often falls short in representing anomalous or extreme events (Lucarini et al., 2016).

DMD was initially proposed as a dimensionality reduction technique that extracts dominant spatio-temporal coherent structures from high-dimensional time-series data. -Motivated by the Occam's razor principle - *The simplest explanation is usually the best one* - our approach tries to extract a single most significant modes that could enlighten the spatial pattern of extreme events. The trade-off between model complexity and accuracy is already explored in (Jovanović et al., 2014). However, the focus there is solely on the sparsity of the model and does not address the reconstruction of extreme values. DMD has significant potential to improve our understanding of the way in which coherent structures in the atmospheric flows evolve and interact (Duke et al., 2012). We frame the problem as an optimization task that places greater emphasis on the optimisation problem designed to preferentially penalise errors at extreme events.

DMD is a practical implementation of the *Koopman theory* in the context of data-driven analysis. Koopman theory focuses on the *Koopman operator*, an infinite-dimensional linear operator that evolves observables over time. It provides a way to analyse the behaviour of nonlinear systems using linear methods - a powerful tool for understanding complex dynamics (Koopman, 1931; Gaspard, 1998).

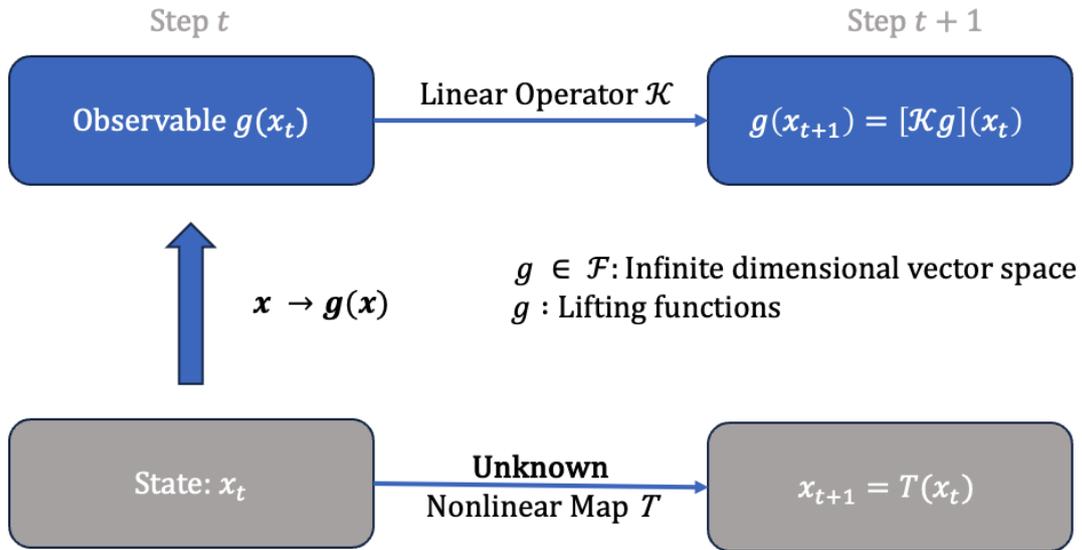


Figure 1. Sketch of the Koopman Operator Theory. Adapted from Shi et al. (2024)

Definition 1. The **Koopman operator** is a linear operator that describes the evolution of observables in a dynamical system. Let $\mathbf{x}_{t+1} = f(\mathbf{x}_t)$ define a discrete-time dynamical system, where $\mathbf{x}_t \in \mathbb{R}^n$ represents the state at time t and f is the state transition map.

An **observable** is a function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ that measures some property of the system state. The Koopman operator, \mathcal{K} , acts on the observable g to describe its evolution:

$$(\mathcal{K}g)(\mathbf{x}) = g(f(\mathbf{x})),$$

or equivalently, in terms of time:

$$g(\mathbf{x}_{t+1}) = (\mathcal{K}g)(\mathbf{x}_t).$$

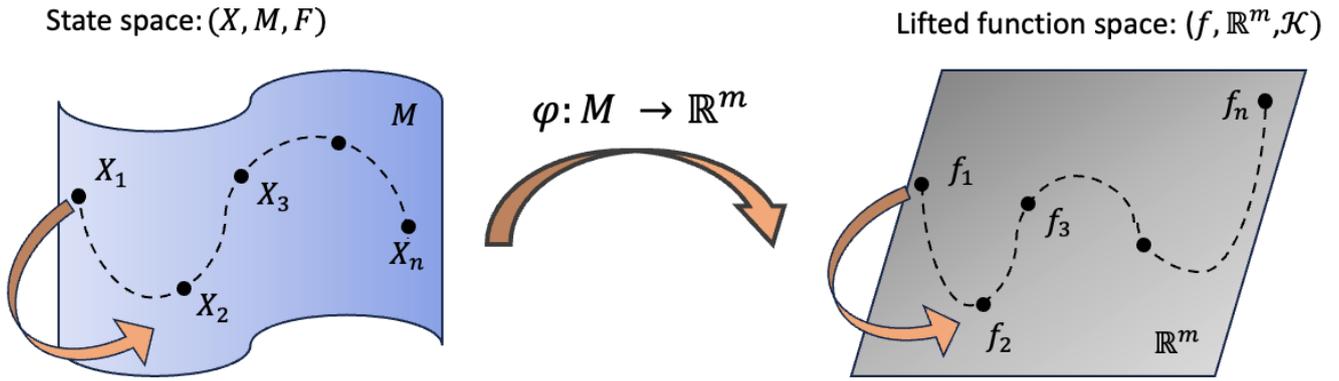


Figure 2. Figure adapted from Wang et al. (2023), demonstrating uplifting.

In other words, the Koopman operator is a linear operator that governs the evolution of scalar functions (often referred to as observables) along trajectories of a given nonlinear dynamical system. Each of the individual measurements may be expanded in terms of the eigenfunction-eigenfunctions $\varphi_j(\mathbf{x})$, which provide a basis for a Hilbert space

$$g(\mathbf{x}_t) = \sum_{j=1}^{\infty} v_{ij} \varphi_j(\mathbf{x}) \quad (1)$$

60

where $\mathbf{v}_j = (v_{1j}, \dots, v_{lj}, \dots)^\top$ is the j -th-th Koopman mode associated with the eigenfunction φ_j . These Koopman modes are coherent *spatial* modes that behave linearly with the same temporal dynamics and are known as dynamic modes in DMD. Now, it is possible to represent the dynamics of the measurements g as follows:

$$g(\mathbf{x}_t) = \mathcal{K}_{\Delta t}^t g(\mathbf{x}_0) = \mathcal{K}_{\Delta t}^t \sum_{j=1}^{\infty} v_{ij} \varphi_j(\mathbf{x}_0) = \sum_{j=1}^{\infty} \mathcal{K}_{\Delta t}^t v_{ij} \varphi_j(\mathbf{x}_0) = \sum_{j=1}^{\infty} \lambda_j^t v_{ij} \varphi_j(\mathbf{x}_0) \quad (2)$$

65 where $\mathcal{K}_{\Delta t}^t$ is Koopman operator applied t times and Δt is the timestep. The triple $(\lambda_j, \varphi_j, \mathbf{v}_j)_{j=0}^{\infty}$ is known as *Koopman mode decomposition* (Mezić, 2005).

The connection between Koopman mode decomposition and Dynamic Mode Decomposition (DMD) was established in (Rowley et al., 2009b). DMD is an algorithm for approximating the Koopman operator (Mezić, 2013; Rowley et al., 2009a), yielding a triple: (i) DMD eigenvalues approximate Koopman eigenvalues λ_j , (ii) DMD modes approximate Koopman modes \mathbf{v}_j (denoted ϕ_j in DMD), and (iii) DMD mode amplitudes approximate Koopman eigenfunctions evaluated at the initial condition, $\varphi_j(\mathbf{x}_0)$ (denoted b_j in DMD).

The main research question is: Can we find the spatial modes that are responsible for an extreme event? Can we find the spatial (DMD) modes that are responsible for the extreme events?

When modeling-modelling dynamical systems, dimensionality reduction algorithms as their name indicates, reduce the dimensionality, i.e. the complexity of the system replacing it with simpler simplify complex systems by replacing them with

75

lower-dimensional models. The common challenge when reconstructing (1), is the question how many modes are needed. This is the balancing act between the simplicity and correctness of the model. ~~By adding less modes we get the signal that is~~ Using fewer modes yields a *smoothed* reconstruction ~~compared to the original signal. By smooth we mean~~, effectively averaging the signal over time ~~omitting the and omitting~~ extremes. Adding more modes improves the reconstruction (e.g., reducing the mean squared error), but it also makes the reconstruction less smooth and causes the extreme values to become more pronounced. In other words, by adding more modes we approximate extremes better.

The questions arise - ~~Can we identify those modes that are responsible for better reconstruction of the extremes? If yes, can we physically interpret these coherent spatial patterns called modes?~~ Can we identify those modes that are responsible for better reconstruction of the extremes? If yes, can we physically interpret these coherent spatial patterns called modes? In spectral analysis, the modes with the highest energy are typically considered. But do the modes with lower energy play a role in extreme events, or are they rightfully dismissed as mere noise?

Probably the most *famous* and widely used method in the dimensionality reduction of physical systems is the Proper Orthogonal Decomposition (POD), also known by other names such as Principal Components Analysis (PCA), the Karhunen-Loeve (KL) decomposition, or empirical orthogonal functions (EOF) (Lorenz, 1956). In climate science EOFs are essentially the result of applying PCA to meteorological data. Consequently, EOF analysis is mathematically equivalent to PCA, and by extension, to POD and Singular Value Decomposition (SVD) (H. Tu et al., 2014). In many nonlinear dynamical systems, snapshot data (measurements) often reveal low-dimensional characteristics, where most of the variance or energy is concentrated in a few modes obtained through SVD. The modes are spatial fields that often identify coherent structures in the flow. However, to the best of our knowledge, all these methods focus on the average behaviour of the nonlinear dynamical system at hand.

Whether we are talking about the numerical method of PCA in statistics (Hotelling, 1933) or POD in fluid mechanics (Berkooz et al., 1993), they both rely on SVD, where the decomposition results are orthogonal vectors that are used for the reconstruction of the original signal. However, the assumption of orthogonality is omitted in the DMD, but the rules of reconstruction stay the same. By omitting the orthogonality, the DMD modes can also be more physically meaningful.

Originally, DMD has been derived from POD (Holmes, 2012; Schmid, 2010; Schmid et al., 2012). However, as already mentioned, POD modes are orthogonal, DMD are not. Additionally, DMD modes are dynamically invariant, POD are not. While POD is based entirely on spatial correlation and energy, DMD adds the temporal information as well. DMD performs a modal decomposition where each mode consists of spatially correlated structures that can be related to certain oscillatory evolution in time. This is a result of the fact that POD solves the eigenvalue problem for the covariance matrix of the data, while DMD employs a time-shifted cross-correlation matrix capturing the linear dependence of the snapshots at the next time step on the snapshots at the current time step (Rowley et al., 2009b; Zhang et al., 2014; Smith et al., 2005).

One of the primary assumptions is that only a few key terms in equation (1) govern the dynamics of a system, resulting in sparse equations within the extensive space of potential functions. Therefore we apply the sparse regression to identify the minimal and most effective terms that accurately capture the underlying dynamics of observables. The resulting model is then parsimonious, striking a balance between complexity and descriptive power, while avoiding overfitting. A similar approach is

110 taken in (Jovanović et al., 2014) where the extension of DMD is developed also with a focus on sparsity of the modes. However, unlike in our extension, the focus there is not on the extremes, but rather on the average signal behaviour.

Our method ~~holds particular promise for addressing~~ shows strong potential for analysing extreme events in the climate system, ~~offering substantial potential for applications across engineering, the physical sciences, and the~~ with applications extending to engineering, physical, and biological sciences. In climate science, DMD ~~can appear under the name of Linear Inverse~~ Modelling is closely related to Linear Inverse Modeling (LIM). Although the algorithms differ, they become equivalent under ~~these three conditions- i) the three conditions:~~ (i) LIM is performed in the space of EOF coefficients (rather than in physical space), ii) the, (ii) EOFs are computed from snapshots alone (no external forcing) (H. Tu et al., 2014) and solely from snapshots without external forcing (H. Tu et al., 2014), and (iii) LIM assumes no stochastic noise.

2 Methodology

120 For better understanding we provide the following illustration of the DMD decomposition (Fig. 3). The ~~input to the method~~ consists method takes as input a sequence of snapshots, ~~with each snapshot each~~ representing the atmospheric state at a specific time step ~~in our experiments~~. The resulting ~~triplet of the DMD mentioned before (DMD triplet - modes, amplitudes, evolution)~~ is well depicted and temporal evolution - is shown schematically here.

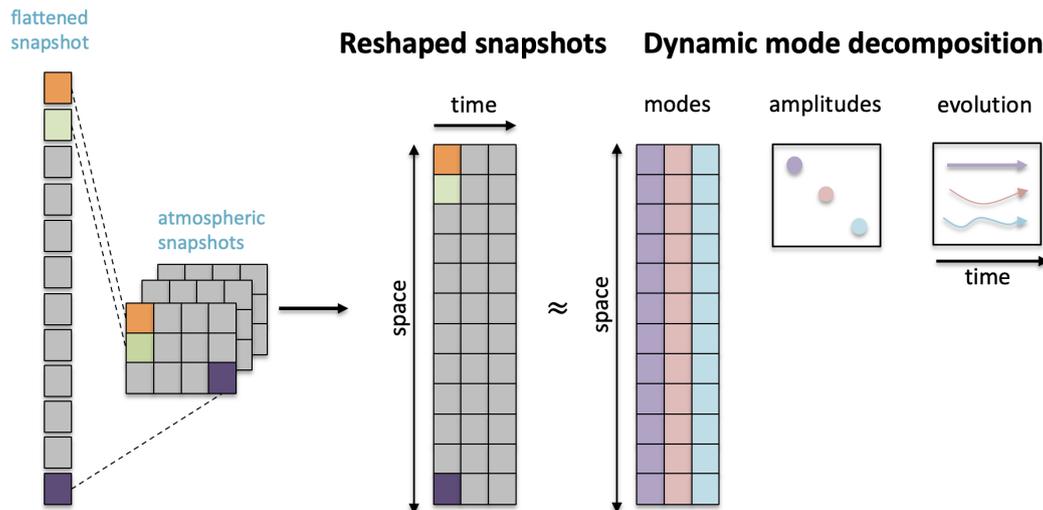


Figure 3. Illustration of the DMD decomposition adapted from Erichson et al., J. Real-Time Image Process., 2016.

125 In this work we will focus on the **amplitudes** matrix and its role in explaining extreme events. The **modes** represent distinct spatial patterns that grow or decay with different frequency over time. The information about growth and decay is captured with the matrix **evolution**. Modes are ordered by their frequencies, with the first modes representing the dominant **behaviors**

behaviours (Smith et al., 2005; Rowley et al., 2009b; Zhang et al., 2014). Each mode has a corresponding amplitude, that holds the information about the energy, i.e. the overall importance for the reconstruction.

130 2.1 Discretisation

In theory, we are interested in the dynamical system and its evolution:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, t; \mu) \quad (3)$$

with a high-dimensional state $\mathbf{x}(t) \in \mathbb{R}^m$. The dynamics $f(\mathbf{x}, t; \mu)$ is unknown, only the observations of the system state can be used to approximate the dynamics and predict the future state. In practice, we are analysing its discrete-time flow map

$$135 \quad \mathbf{x}_{k+1} = F(\mathbf{x}_k) \quad (4)$$

We search for the operator A that linearises the dynamics in form of:

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x} \quad (5)$$

that has a closed-form solution:

$$\mathbf{x}(t_0 + t) = e^{At}\mathbf{x}(t_0) \quad (6)$$

140 The solution to this system may be expressed simply in terms of eigenvalues ~~λ_k and eigenvectors ϕ_k~~ λ_j and eigenvectors ϕ_j of this discrete map A .

$$x_{k+1} = \sum_{j=1}^K \phi_j \lambda_j^{k+1} b_j \quad (7)$$

where $b_j \in \mathbb{C}$, also called amplitude, contains the weights of the modes, the eigenvalues $\lambda_j \in \mathbb{C}$ represent the temporal behaviour of the system and the eigenvectors ~~$\phi_j \in \mathbb{C}^n$~~ $\phi_j \in \mathbb{C}^n$ are the spatial modes. Equation (7) is a mathematical formulation of Fig. 3.

~~Depending on the dynamics, we might be able to perform~~ In theory, such a decomposition ~~exactly can be exact~~ using finitely many modes; ~~in real and realistic flows, this is highly unlikely,~~ but in realistic flows this is rarely achievable. Instead, we aim to achieve the decomposition only approximately, by retaining a relatively small number of modes K .

2.2 Regression

150 The DMD method produces a low-rank decomposition of the matrix A by optimising the fit to the trajectory x_k

$$\|x_{k+1} - \mathbf{A}x_k\|_2 \quad (8)$$

where x_k and x_{k+1} represent measurements taken at two consecutive timestamps, separated by time interval Δt .

This linearisation ~~holds only locally as we are finding optimal local linear approximations of snapshots, where snapshots are states of the systems sampled in time~~ is valid only locally, as it represents the best local linear approximation of the system's time-evolving states (snapshots).

Since we want to minimise the error across all snapshots, we reshape all m ~~snapshots~~ snapshots into a high-dimensional column vectors and arrange them in two large matrices:

$$\mathbf{X} = \begin{bmatrix} | & | & | & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{m-1} \\ | & | & | & | \end{bmatrix} \quad (9)$$

$$\mathbf{X}' = \begin{bmatrix} | & | & | & | \\ \mathbf{x}_2 & \mathbf{x}_3 & \dots & \mathbf{x}_m \\ | & | & | & | \end{bmatrix} \quad (10)$$

and approximate the linear operator A that models the trajectory $x(t)$ by formulating

$$\mathbf{X}' \approx \mathbf{A}\mathbf{X} \quad (11)$$

Putting it all together we have the following definition.

Definition 2 (DMD). *Suppose we have a dynamical system as in (3) and two sets of data (snapshots) \mathbf{X} and \mathbf{X}' as in (9) and (10) so that $x'_k = F(x_k)$ is the map in (4) corresponding to the map (3) for time Δt . DMD computes the leading eigendecomposition of the best-fit linear operator \mathbf{A} relating the data $\mathbf{X}' \approx \mathbf{A}\mathbf{X}$ (H. Tu et al., 2014).*

Mathematically, the best-fit operator \mathbf{A} is then defined as

$$\mathbf{A} = \arg \min_{\mathbf{A}} \|\mathbf{X}' - \mathbf{A}\mathbf{X}\| = \mathbf{X}'\mathbf{X}^\dagger \quad (12)$$

where \dagger denotes the Moore-Penrose pseudo-inverse of a matrix.

The validity of the approximation in (12) is supported by the following theorem, which relies on the concept of linear consistency. ~~Let~~ But let us first introduce the definition of the linear consistency.

Definition 3. *Two matrices $\mathbf{X} \in \mathbb{C}^{n \times m}$ and $\mathbf{X}' \in \mathbb{C}^{n \times m}$ are said to be linearly consistent if for every vector $\mathbf{c} \in \mathbb{C}^m$, the condition $\mathbf{X}\mathbf{c} = 0$ implies $\mathbf{X}'\mathbf{c} = 0$.*

Theorem 1 (Tu et al., 2014, (H. Tu et al., 2014)). *If we define $\mathbf{A} = \mathbf{X}'\mathbf{X}^\dagger$. Then $\mathbf{X}' = \mathbf{A}\mathbf{X}$ if and only if \mathbf{X}' and \mathbf{X} are linearly consistent.*

We conclude that by inverting the usual focus of the method ~~—from modeling from modelling~~ the system's average ~~behavior~~ behaviour to emphasizing its exceptional dynamics ~~—and by carefully refining the algorithm accordingly,—~~ Additionally we

refine the DMD algorithm accordingly so that extreme events can be modeled more accurately across several common metrics. Moreover, this approach enables the extraction of spatial patterns that contribute to the reconstruction of extremes, along with insights into their temporal evolution.

180 Even though in this work we focus on the diagnostic application, the method could be easily used for future state prediction. ~~This extension is planned for the future work.~~ Diagnosing the evolution in time of extreme situations, one could detect reoccurring patterns and therefore predict the upcoming extremes. This extension is planned for the future work.

2.3 DMD Algorithm

As ~~mentioned before~~previously noted, the DMD algorithm ~~produces~~identifies the best-fit linear operator that ~~relates the two~~
185 ~~snapshot matrices in time~~advances the system from one snapshot matrix to the next in time.

1. Compute the singular value decomposition of \mathbf{X} (defined in (9)), performing a low-rank truncation at the same time

$$\mathbf{X} \approx \tilde{U} \tilde{\Sigma} \tilde{V}^* \quad (13)$$

where $*$ denotes the conjugate transpose, $\tilde{U} \in \mathbb{C}^{n \times r}$, $\tilde{\Sigma} \in \mathbb{C}^{r \times r}$, and $\tilde{V} \in \mathbb{C}^{m \times r}$, and $r \leq m$ denotes either the exact or the approximate rank of the data matrix \mathbf{X}

- 190 2. Project the full matrix A onto \tilde{U} and calculate \tilde{A}

$$\tilde{A} = \tilde{U}^* A \tilde{U} = \tilde{U}^* X' \tilde{V} \tilde{\Sigma}^{-1} \quad (14)$$

3. Compute the eigenvalues λ_i (diagonal of matrix Λ) and eigenvectors w_i (columns of W) of \tilde{A} . These eigenvalues and eigenvectors also correspond to the ones of the full matrix A .

$$\tilde{A} W = W \Lambda \quad (15)$$

- 195 4. The high-dimensional DMD modes Φ are reconstructed using the eigenvectors W and the time-shifted snapshot matrix X' . These DMD modes are eigenvectors of the high-dimensional matrix A corresponding to the eigenvalues in Λ .

$$A \Phi = \Phi \Lambda, \quad \Phi = X' V \Sigma^{-1} W \quad (16)$$

5. Reconstruct the original signal

$$\mathbf{X}_{dmd} = \Phi e^{\Omega t} \mathbf{b}, \quad \mathbf{b} = \Phi^\dagger x_0 \quad (17)$$

200 The angular frequency ω_k associated with each eigenvalue λ_k is calculated $\omega_k = \frac{\ln(\lambda_k)}{\Delta t} b$, though alternative approaches exist. In our experiments, we will further explore these alternatives using ~~optimization~~optimisation techniques.

2.4 Optimisation

Our idea is closely related to the sparse DMD (Rudy et al., 2016), where the aim is to reconstruct the original signal using as few modes as possible. In other words, to find the solution that exhibits the best balance of accuracy and sparsity. This is implemented by adding a regularisation term γ in the regression step (8), i.e. the L^1 penalty.

$$\min_{\mathbf{b}} J(\mathbf{b}) + \gamma \sum_{i=1}^r |b_i| \quad (18)$$

where the objective function $J(b)$ is defined by

$$J(b) := \|X - \Phi D_b V_\mu\|_F^2 = \sqrt{\sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \Phi_{ij} D_{b_{ij}} V_{\mu_{ij}})^2} \quad (19)$$

where Φ denotes the matrix of DMD modes, as defined in Step 4 of the algorithm above, $D_b = \text{diag}\{\mathbf{b}\}$ has the amplitudes on the diagonal and V_μ is the Vandermonde matrix holding information about the angular frequency.

The goal is to control the amplitudes of the vector $\mathbf{b} = \Phi^\dagger x_o$. The vector \mathbf{b} is needed to reconstruct a matrix corresponding to the time evolution of the system. It is used to reconstruct the system's time-evolution matrix and serves as the "starting point" for the dynamical system analysed. The central question of the analysis. A central questions in sparse DMD is: What are the best modes for a system, and how can they be identified? Which modes best represent a system, and how can they be identified? This is a complex, non-trivial challenging problem. The proposed method aims to automate this process by reconstructing the data using as few DMD modes as possible. But the challenge remains — Which modes should be chosen? With - Which modes should be chosen?

In our DMD variation, we refine the previous question: What are the best modes for reconstructing the focus on identifying the modes that best reconstruct the system's extremes of a system, and how can they be identified? To answer this question, we modify the objective function (19) by adding an additional regularisation term that accounts for extreme events. We will call it a penalisation term, since it penalises the reconstructions in which the extremes are poorly represented.

$$J(b) = \|X - \Phi D_b V_\mu\|_F^2 = \sqrt{\sum_i \sum_j (X_{ij} - \Phi_{ij} D_{b_{ij}} V_{\mu_{ij}})^2} + \underbrace{\sqrt{\sum_{i \in M} \sum_{j \in M_i} (x_{ij} - \phi_{ij} D_{b_{ij}} V_{\mu_{ij}})^2}}_{\text{penalisation term}} \quad (20)$$

where M represents a set of temporal indices of extreme events and M_i is a set of spatial indices for a certain extreme event i . In equation (20), both the global and the extreme-set residuals are measured using the Frobenius norm. The second term represents a Frobenius-norm penalty applied to the subset of extreme observations.

Here, the penalisation term suppresses deviations from those. The penalisation term reduces deviations from extreme events. In this way Consequently, the objective function is optimized for the vector b , favoring the amplification of optimizes b to favour extraordinary modes — those responsible for detecting outliers — rather than the, over average ones that approximate average modes that capture typical behaviour.

230 3 Experiments

In our experiments, we ~~utilize~~utilise reanalysis data, as it represents the most accurate approximation to observations currently available. Although reanalyses are model-based reconstructions, they are constrained by observational data. Variables that are directly assimilated into the reanalysis forecast model tend to align more closely with real-world measurements.

For this study, we use the ERA5 reanalysis dataset (Hersbach and Dee, 2016), downloaded on a regular $0.25^\circ \times 0.25^\circ$ grid. Our spatial domain is restricted to Europe, spanning 70°N to 36°N and 9°W to 32°E , yielding a resolution of 17892×31224 grid points. The temporal coverage extends every day over 82 years, from 1940 to 2022. In the following experiments we focus on temperature anomaly in Europe. Although the method can be applied to any atmospheric variable.

The assumption is that the measurements are taken from the unknown dynamics and the research question is: *Are there any coherent patterns of extremes from the hidden dynamics that can be discovered using DMD algorithm?* To answer this question we implement the new variation of a ~~Sparse~~-Dynamic Mode Decomposition (~~SPDMD~~DMD), called *extreme* DMD (defined in (20)).

3.1 Analysis

In our experiments we compare two DMD variations, the *normal* and the *extreme* one, that includes an additional penalisation term. Having both reconstructions, we will ~~analyze which of both~~analyse which of them performs better and which modes are significant in each of the reconstructions.

As previously mentioned, the significance information is encoded in the amplitude vector \mathbf{b} . By comparing the two amplitude vectors, \mathbf{b}_{normal} and $\mathbf{b}_{extremes}$, we can determine which modes carry greater significance and which are less relevant when dealing with extreme events. Once these key modes are identified, the next step is to ~~explore whether a physical interpretation exists.~~investigate their physical interpretation, which we leave for future work.

We extract the *most important mode* in this manner:

$$\text{most important mode index} = \arg \max \{ \mathbf{b}_{\text{extreme}} - \mathbf{b}_{\text{normal}} \} \quad (21)$$

3.2 Persistent Extreme Event Detection and Validation

First let us define how we extract the persistent extreme events.

3.3 ~~Persistent Extreme Event Detection and Validation~~

In our test cases, we focus on heatwaves.

Heatwaves are prolonged periods of unusually hot weather, though their exact definitions vary across the literature (Xu et al., 2016). Typically, heatwaves are defined based on a combination of duration and high-temperature thresholds (Xu et al., 2016; Yin et al., 2022; Casati et al., 2013), often involving percentile-based approaches or specific indices, such as the ETCCDI indices (Vogel et al., 2020). Studies have shown that the choice of threshold influences projected changes in heatwave characteristics (Vogel et al., 2020), and this choice differs based on the regional climate (Perkins-Kirkpatrick and Gibson, 2017).

Therefore, climate extremes such as heatwaves, are not only defined by peak values, they must also persist over a specific time period within a region. In our experiments, we define ~~this persistence as a duration of at least~~ extreme heat events using a 95th percentile threshold and require a minimum duration of three consecutive timesteps above this level.

3.3 Local Maximum Detection

265 To identify extreme persistent events, we first detect a local maximum in a snapshot. We then compare this maximum to the following two snapshots to check if the maximum persists. The condition is that the maximum stays within the neighborhood (e.g. 5x5 pixels).

Let $D(t, x, y)$ represent the data value at time t and spatial location (x, y) . N represent the neighborhood size for initial detection.

270 For each time step t^* , identify (x^*, y^*) as the location of a local maximum :

$$D(t^*, x^*, y^*) = \max_{(i,j) \in N} D(t^*, x^* + i, y^* + j)$$

3.4 Persistence Check Over Previous Time Step

For a local maximum (x^*, y^*) at time t , check persistence over the previous two time steps, $t - 1$ and $t - 2$:

$$D(t - 1, x^*, y^*) = \max_{(i,j) \in N} D(t - 1, x^* + i, y^* + j)$$

$$D(t - 2, x^*, y^*) = \max_{(i,j) \in N} D(t - 2, x^* + i, y^* + j)$$

If **Local Maximum Detection** and **Persistence Check over Previous Time Steps** are satisfied, only then (t^*, x^*, y^*) is considered a persistent maximum. The procedure to find minima works analogously. Importantly, the detection of extreme events
 275 is performed in the transformed modal space, rather than in the original data domain. This approach allows us to identify and isolate the specific dynamic modes that contribute most significantly to rare or extreme ~~behavior~~ behaviour. As a result, it provides a more robust and interpretable basis for identifying outliers, compared to simply searching for large values in the raw data. Moreover, because the modal representation operates in a reduced-dimensional subspace, this approach is also computationally more efficient, making it well-suited for high-dimensional systems.

280 3.5 Extreme Dynamic Mode Decomposition

We modify the *original* DMD by adding the penalisation term, which is then solved by a convex optimization problem that balances the reconstruction error ~~-, and~~ extreme value penalisation, ~~and regularization term~~. We are minimising the objective function defined in (20). We test the method across various ranks, but represent in the following section the results for rank 50.

3.6 Metrics

285 ~~We use 4 different metrics to measure the fitness of the results. They are~~ Reconstruction accuracy is evaluated using four distinct performance metrics:

- **Mean squared error (MSE)** measures the average squared difference between predicted values \hat{y} and the actual (true) values y

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

- **L-infinity norm (L_∞)** (also called the maximum norm or supremum norm) of a vector measures the largest absolute value among its components. In other words, it captures the maximum magnitude of any single element in the vector.

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

- 290
- **The 10-norm** is a specific case of the ~~Lp-norm~~ L_p -norm where $p = 10$. It measures the "length" of a vector by taking the ~~10th~~ 10-th root of the sum of the absolute values of the vector components raised to the 10th power. Higher-order norms (like the 10-norm) raise the absolute values to a high power before summing, which amplifies the effect of large components (i.e., outliers).

$$\|\mathbf{x}\|_{10} = \left(\sum_{i=1}^n |x_i|^{10} \right)^{\frac{1}{10}}$$

- 295
- **Structural Similarity Index Measure (SSIM)** is a perceptual metric usually used to measure the similarity between two images (or signals) by comparing their luminance, contrast, and structural information. Unlike traditional error metrics (like MSE), SSIM aims to model the human visual system's sensitivity to changes in structural information, making it better suited for assessing perceived image quality.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

- 300
- where:

- μ_x and μ_y are the mean intensities of image patches x and y , respectively:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i, \quad \mu_y = \frac{1}{N} \sum_{i=1}^N y_i$$

- σ_x^2 and σ_y^2 are the variances of x and y , representing contrast:

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2, \quad \sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2$$

- 305
- σ_{xy} is the covariance between x and y , capturing structural similarity:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$$

– C_1 and C_2 are small constants to stabilize the division, defined as:

$$C_1 = (K_1 L)^2, \quad C_2 = (K_2 L)^2$$

where L is the dynamic range of pixel values (e.g., 255 for 8-bit images), and $K_1, K_2 \ll 1$.

310 3.7 Results

To facilitate interpretation of the following figures, we briefly recall the meaning of amplitudes and modes in the context of Dynamic Mode Decomposition (DMD). Each DMD mode represents a spatially coherent structure that evolves over time with a specific frequency and growth or decay rate. The associated amplitude quantifies the importance of that mode in reconstructing the observed system dynamics. In the proposed *extreme* DMD framework, the amplitude estimation is modified by a penalisation term that prioritises accurate reconstruction of extreme events. Therefore, differences in amplitudes between the Normal and *extreme* DMD indicate which spatiotemporal structures gain or lose influence when the method focuses on extremes rather than average behaviour. This interpretation underpins Figs. 5, 9, and 16, which highlight how the relative weighting of modes shifts in response to the extreme-event constraint.

3.7.1 Heatwave 2003

We now examine in detail the well-known heatwave that struck France in August 2003, one of the most extreme and widely studied heat events in recent history (García-Herrera et al., 2010). This heatwave was characterized by persistently high temperatures, leading to severe impacts on human health, agriculture, and infrastructure. The spatial extent of the heatwave during its peak in early to mid-August 2003 is illustrated in

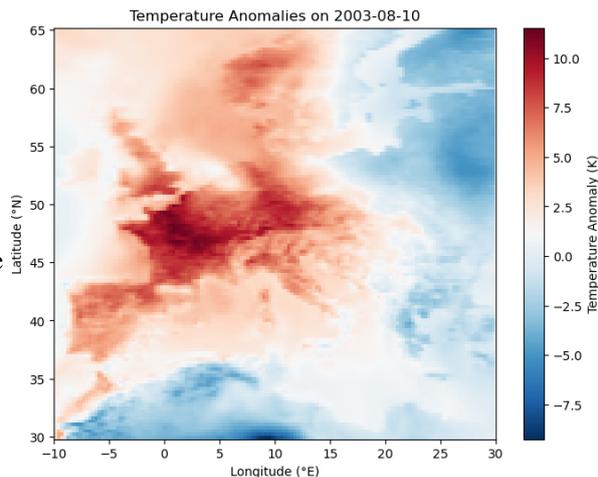


Figure 4. Heatwave of 2003: Red shading indicates increased temperature anomalies (°C) over Europe on August 03, 2003.

320 To clarify the methodology used in our experiments, in Fig. 5 we present a schematic analogue to Fig. 3, but constructed using our dataset corresponding to the 2003 heatwave. This illustration serves to provide a clearer understanding of the experimental setup and the specific data processing steps applied in our analysis.

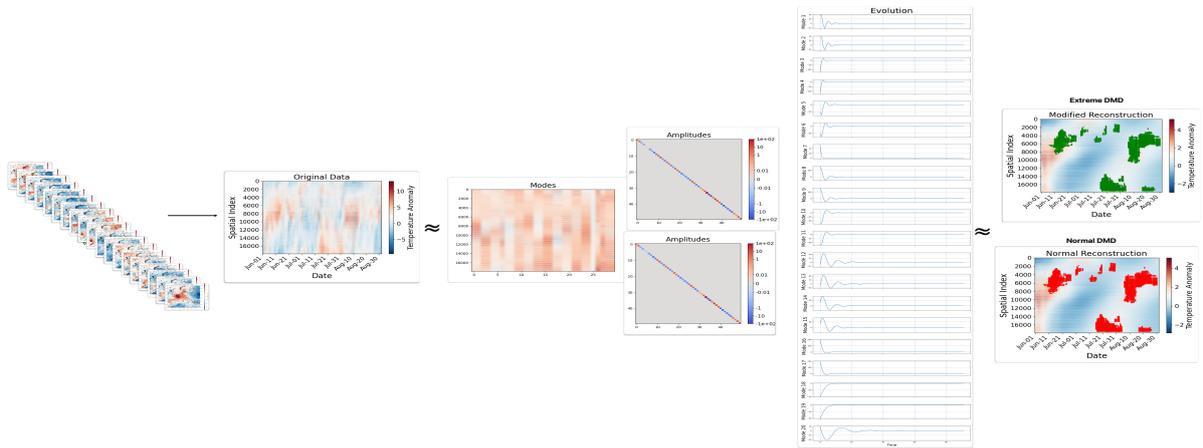


Figure 5. Comparison of normal and *extreme* Dynamic Mode Decomposition (DMD) applied to the summer 2003 data, following dataset. Both methods follow the same procedure as in Fig. 3. The figure compares 3 but differ in the amplitude estimation step, as shown by the two DMD variants distinct amplitude matrices in parallel the scheme. The temporal evolutions of the spatial patterns remains identical for both approaches (see Evolution plots in the scheme). The combination of modes, amplitudes, and evolution produces the reconstructed surface temperature anomalies shown in the rightmost panels: *Normal* the upper panel corresponds to the Normal DMD reconstruction, and *Extreme* the lower panel to the *extreme* DMD. The primary difference lies in *extreme* DMD, the amplitudes are optimised with an additional penalisation term that emphasizes the accurate reconstruction of extreme temperature values. Differences between the two reconstructions highlight how weighting extremes alters the spatial distribution and magnitude of the reconstructed anomalies, which lead leading to slightly different reconstructions an improved representation of the original data 2003 European heatwave.

Which reconstruction is better?

To show that the *extreme* DMD results indeed with better reconstruction, we show the following resulting reconstructions.

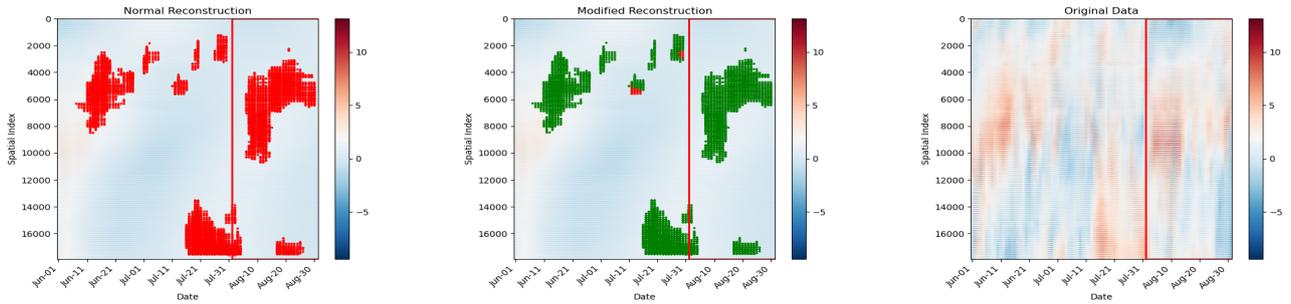


Figure 6. Reconstruction comparison Comparison of **Heatwave 2003**. All three plots have x-axis as a reconstructions for the 2003 heatwave. Each plot shows time dimension and y-axis as (x-axis) versus space dimension. They are (y-axis), corresponding to the flattened snapshots illustrated in the Fig. 3. The plotted markers (red Red and green) represent markers denote the occurrence of the extreme event events. The left plot represents the reconstruction using the normal (a) Normal DMD. The middle plot represents the reconstruction using the, (b) extreme DMD reconstruction, where green color indicates the highlights indicate extremes that are better reconstructed captured (having the smaller lower MSE) compared to the normal normal DMD case, i.e. left plot. Whereas the right plot serves and (c) original anomaly field serving as a reference plot, representing the original values of the anomalies. Red The red frame indicates snapshots of only marks the August period, when most of the extreme events occurred. Both plots are results of All reconstructions follow the scheme shown in Fig. 5-5.

325 In Fig. 6, the left panel illustrates the reconstruction obtained using the standard (normal) Dynamic Mode Decomposition (DMD) without a penalization-penalisation term, while the middle panel presents the results of the modified extreme DMD approach, which incorporates a penalization-penalisation term. The rightmost panel displays the original data for reference. The dots in the plots denote extreme anomalies, with those in the middle panel highlighted in green when the reconstructed values closer match the original extreme events, as indicated by a lower mean squared error (MSE) and other metrics (see Fig. 330 7. and Fig. 8). We focus on August (indicated with the red rectangular in Fig 6.), the peak of the heatwave period.

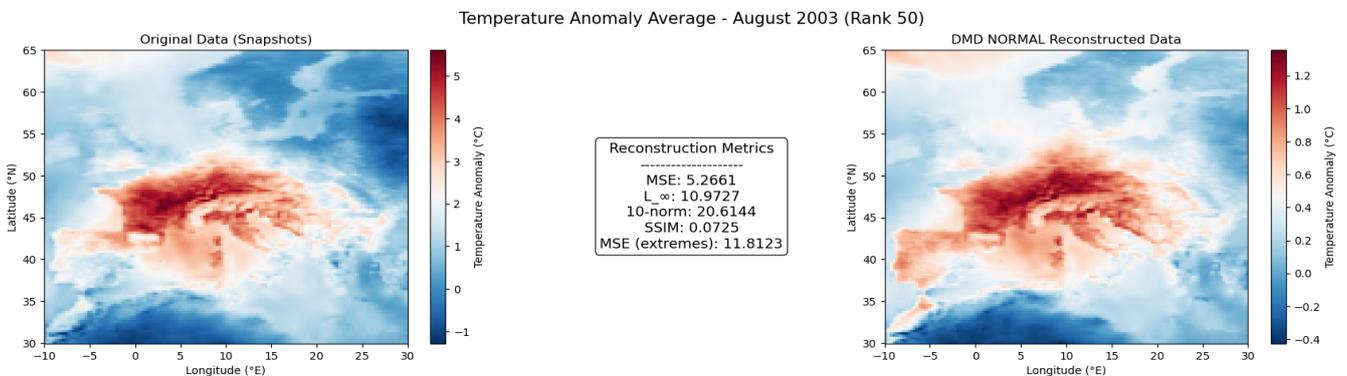


Figure 7. The left panel shows the average temperature anomaly in August based on the original ERA5 data (reference), while the right panel shows the corresponding reconstruction using the normal DMD method. The middle panel presents various metrics that evaluate the quality of the reconstruction.

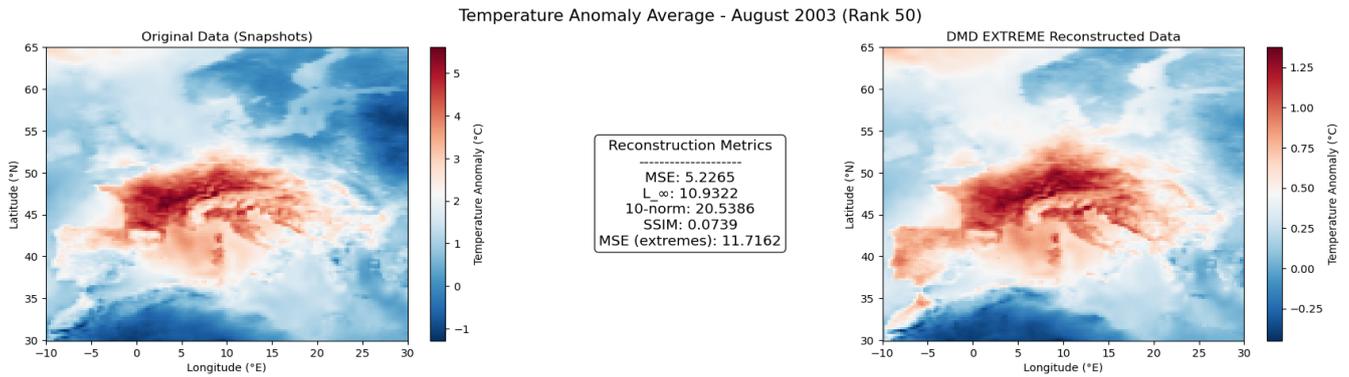


Figure 8. The left panel shows the average temperature anomaly in August based on the original ERA5 data (reference), while the right panel shows the corresponding reconstruction using the novel *extreme* DMD method. The middle panel presents various metrics that evaluate the quality of the reconstruction.

By ~~analyzing~~ analysing and comparing various metrics presented in Fig. 7. and Fig. 8., we observe that the differences are subtle but systematic and aligned with our goal of capturing extremes. The proposed method - *extreme DMD* (results in Fig. 8) - offers a more accurate representation of extreme values.

335 Statistical comparison between the *extreme* DMD and standard DMD reconstructions at extreme spatiotemporal points (MSE (extremes)) showed a significant improvement for the *extreme* DMD ($t = 90.150$, $p < 0.001$). The effect size (Cohen's $d = 0.558$) indicates a moderate enhancement, with a 95% bootstrap confidence interval for the mean MSE reduction of $[0.094103, 0.098255]$. These results confirm the superior reconstruction accuracy of the *extreme* DMD method during extreme events.

Which modes have more significance and which less in each of the reconstructions?

340 After running two different ~~optimisation~~ optimization problems: without penalising the extreme events (minimising (19)) and once with penalising the extreme events (minimising (20)), we compare the resulting optimal solutions - the amplitudes b 's - and we search for the biggest absolute difference.

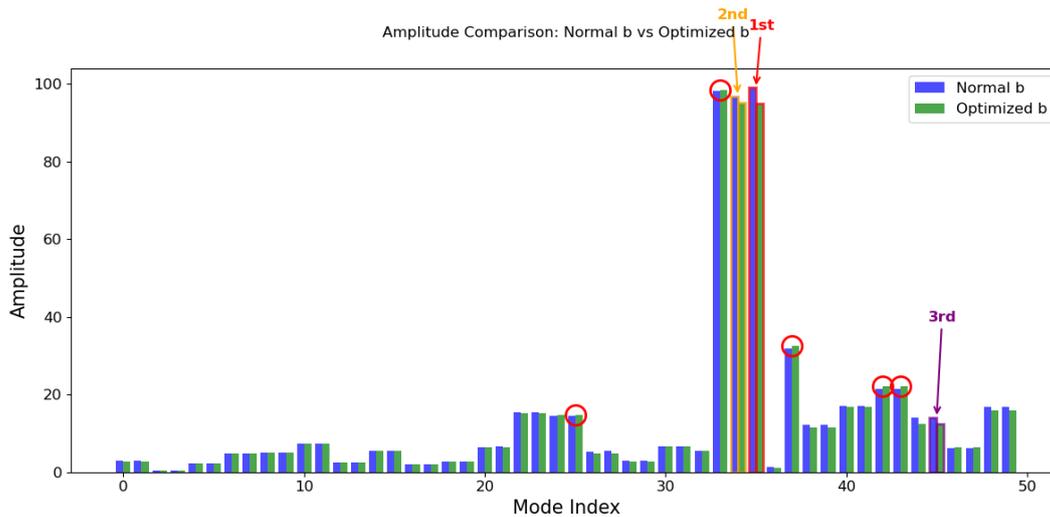


Figure 9. Amplitude differences Comparison of mode amplitudes between the normal and extreme Dynamic Mode Decomposition (DMD) for the 2003 heatwave event. Each bar represents the amplitude of a single mode, quantifying its contribution to the reconstructed temperature field. The largest absolute amplitude difference between the two methods is highlighted with a red frame, the second largest with an orange frame, and the third with a purple frame. Additional modes marked with red circles indicate modes that deviate from the general trend smaller but notable deviations, exhibiting where the extreme DMD assigns higher amplitude values in than the Extreme normal DMD compared. These highlighted differences identify which modes are most affected by the extreme-event penalisation term and thus contribute more strongly to the Normal DMD reconstruction of extreme temperature anomalies.

In this experiment (with rank 50), the most important mode related to extreme events—, defined by the highest amplitude difference—is mode 35—, is mode 35 (see Fig. 9). We extract this mode along with its temporal dynamics and present them in Fig. 10.

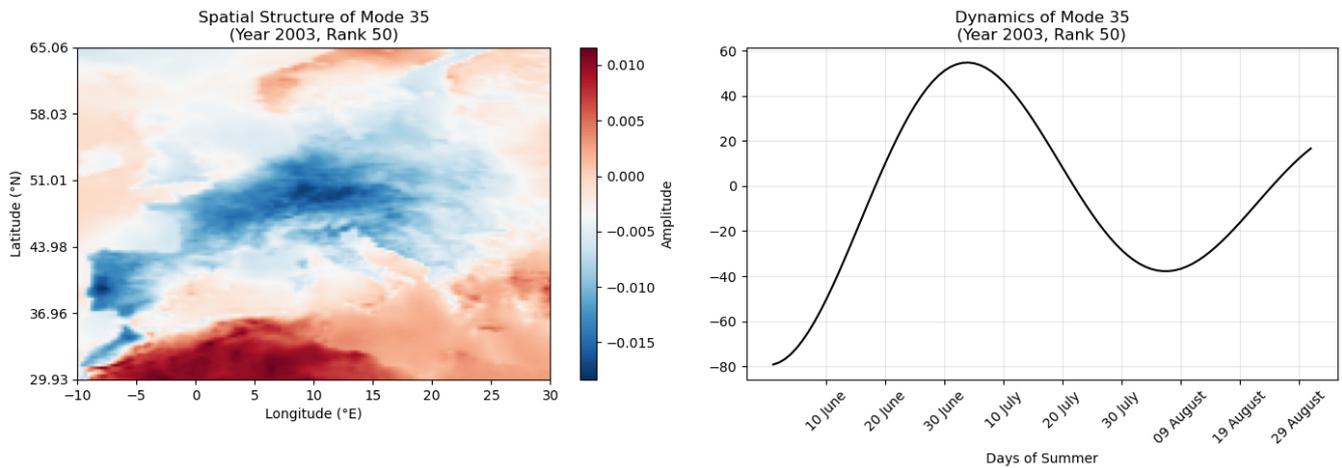


Figure 10. Spatial pattern and the dynamics of the *1st most important mode*.

345 **Based on** Fig. 10 illustrates that the pronounced anomaly contrast between Europe and northern Africa is more influential in shaping the average **behavior-behaviour** of temperature anomalies than the extreme events. The temporal evolution of this mode reaches its maximum influence in late June, as evident on the right side of the panel.

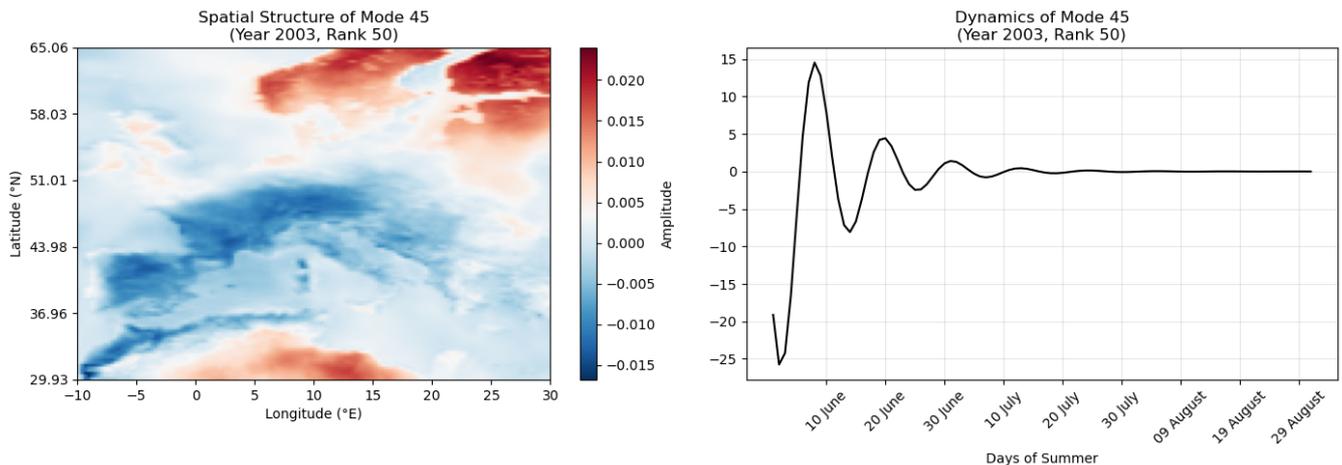


Figure 11. Spatial pattern and dynamics of the mode **45 - mode** exhibiting the greatest increase in **Optimized-optimised b** ($b_{extreme}$) value relative to the **Normal-normal b** (b_{normal}) value (red circles in the Fig. 9).

A spatial pattern exhibiting higher amplitude values in the extreme DMD reconstruction (unlike the others) shown in Fig. 11 suggests that the pronounced contrast in anomalies between **Eastern-Northeastern** and Western Europe played one of the key
 350 role in driving the 2003 heatwave in France.

3.7.2 Heatwave 2010

In Northern Germany, the heatwave in 2010 led to unusually high temperatures, with several regions experiencing prolonged periods above 35°C (Barriopedro et al., 2011). The extreme heat contributed to drought conditions, affecting agriculture and low water levels in rivers such as the Elbe and Weser. Additionally, urban areas like Hamburg and Bremen recorded significantly above-average temperatures, causing discomfort and increasing energy demand for cooling. While the heatwave in Germany was not as intense as in Russia, it still had noticeable impacts on public health and infrastructure. Here, we present the temperature anomalies that affected Northern Europe in mid-July.

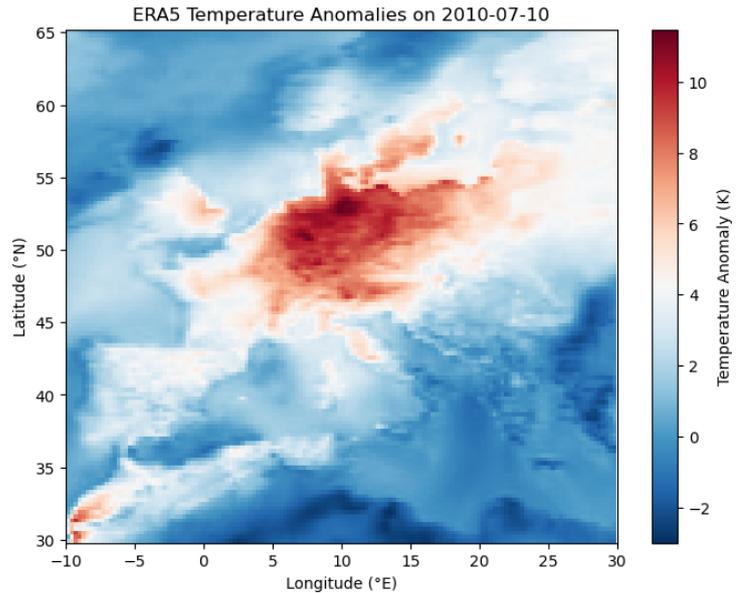


Figure 12. Heatwave of 2010: Red shading indicates the Spatial distribution of temperature anomaly-anomalies (°C) over Europe on July 10, 2010.

Which reconstruction is better?

To evaluate the results, we conduct the reconstruction comparison using the same approach as in the previous experiment for the 2003 heatwave. The plotted dots denote extreme anomalies, with green highlighting regions where the reconstruction aligns more closely with the original extreme values.

The results clearly demonstrate that the *extreme* DMD approach leads to a more accurate representation of extreme temperature anomalies. This improvement is quantified by a lower mean squared error (MSE), both for the overall reconstruction and specifically for the extreme anomalies. Notably, all major extreme values observed during the 2010 heatwave are better captured using the modified DMD approach, further reinforcing the effectiveness of incorporating penalization-penalisation in detecting and reconstructing extreme climate events. This behavior-behaviour is further quantified by other metrics shown in Fig. 14. and Fig. 15..

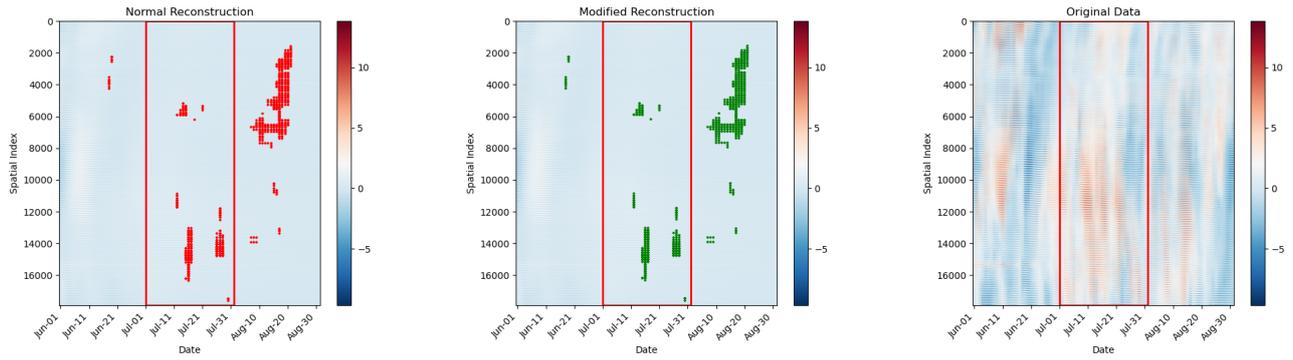


Figure 13. Reconstruction comparison Heatwave 2010. All three plots have x-axis as a Comparison of reconstructions for the 2010 heatwave. Each plot shows time dimension and y-axis as (x-axis) versus space dimension. They are (y-axis), corresponding to the flattened snapshots illustrated in the Fig. 3. The plotted markers (red Red and green) represent markers indicate the occurrence of the extreme event events. The left plot represents the reconstruction using the normal (a) Normal DMD . The middle plot represents the reconstruction using the (b) extreme DMD reconstruction, where green color indicates the highlights indicate extremes that are better reconstructed captured (having the smaller lower MSE) compared to the normal normal DMD, i.e. left plot. Whereas the right plot serves and (c) original anomaly field as a reference plot, representing the original values of the anomalies. The red frame marks the July snapshots, corresponding to the period, when extreme events had the greatest impact.

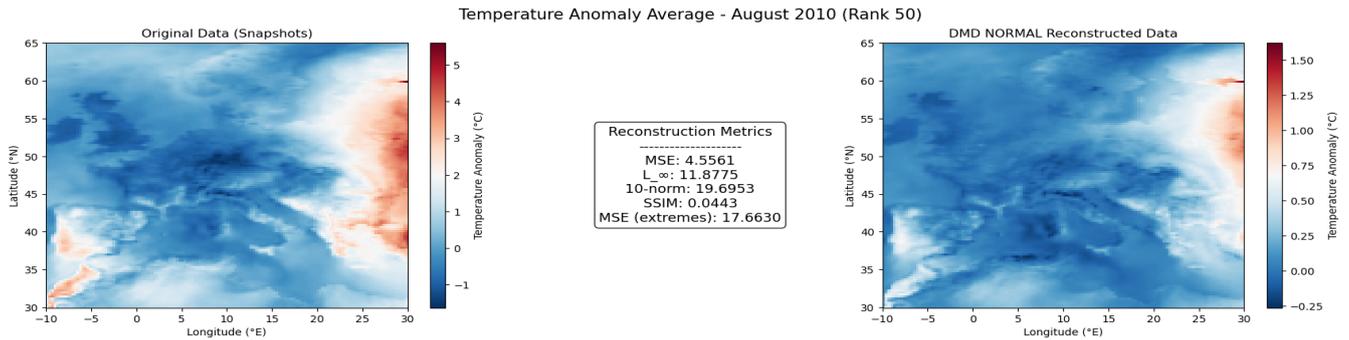


Figure 14. The left panel shows the average temperature anomaly in July based on the original ERA5 data (reference), while the right panel shows the corresponding reconstruction using the normal DMD method. The middle panel presents various metrics that evaluate the quality of the reconstruction.

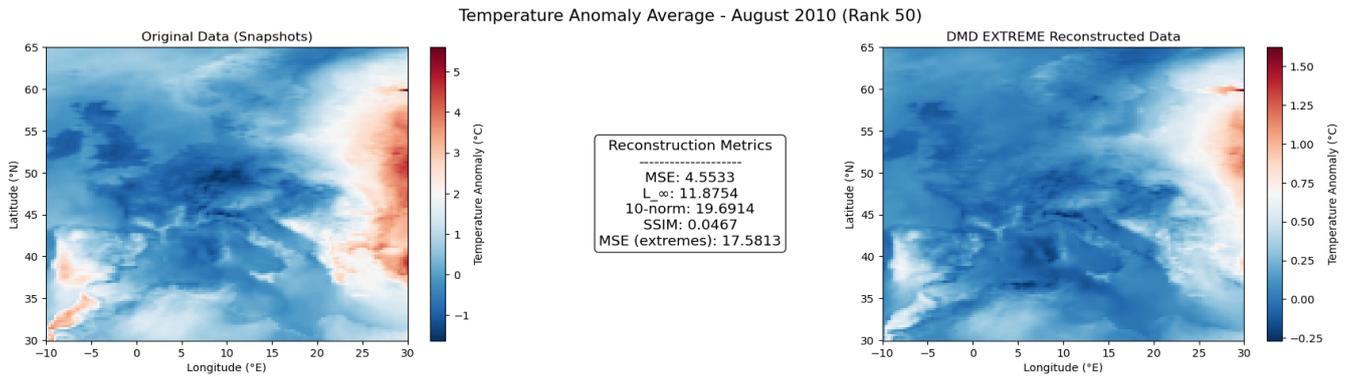


Figure 15. The left panel shows the average temperature anomaly in July based on the original ERA5 data (reference), while the right panel shows the corresponding reconstruction using the novel *extreme* DMD method. The middle panel presents various metrics that evaluate the quality of the reconstruction.

Consistent with earlier findings, the differences in performance metrics presented in Fig. 14. and Fig. 15. are subtle yet systematic, and once again align with our objective of enhancing the representation of extreme events. The proposed method -
 365 *Extreme-DMDE extreme DMD* (results shown in Fig. 15) - demonstrates improved accuracy in capturing extreme values.

Statistical evaluation further supports this improvement. A paired *t*-test revealed a highly significant difference between the two reconstructions ($t = 29.82$, $p < 0.001$), with a large effect size (Cohen's $d = 0.83$). The 95% bootstrap confidence interval for the mean improvement ranged from 0.076 to 0.087, confirming that *extreme DMD* consistently outperforms the standard DMD approach in reproducing extreme temperature anomalies.

370 **Which modes have more significance and which less in each of the reconstructions?**

Again we compare the amplitude vectors \mathbf{b} 's which are optimisation results with the penalisation term (minimising (20)) and without (minimising (19)).

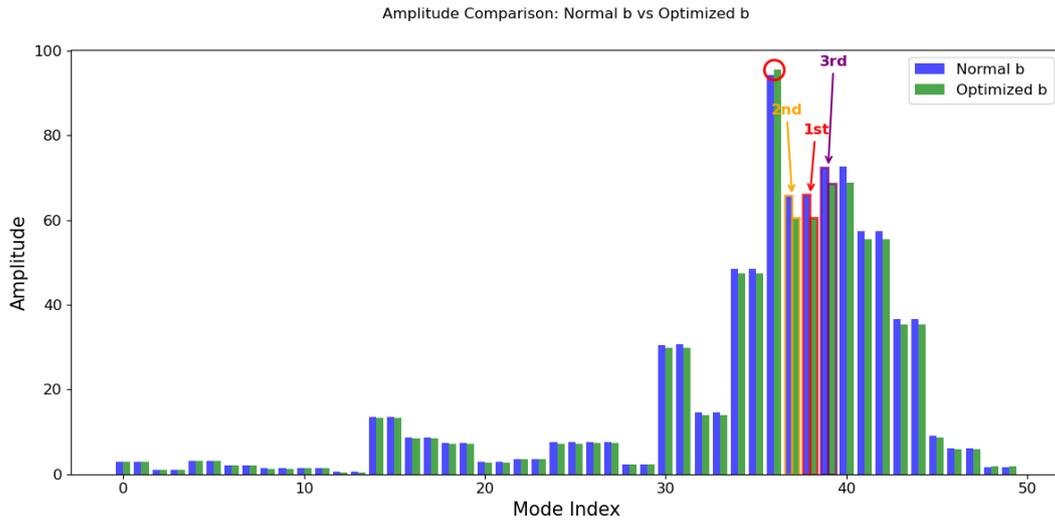


Figure 16. ~~Amplitude differences~~ Comparison of mode amplitudes between ~~modes~~ the *normal* and *extreme* Dynamic Mode Decomposition (DMD) for the 2010 heatwave event. Each bar represents the amplitude of a mode, indicating its relative contribution to the reconstructed temperature field. The largest absolute amplitude difference between the two methods is highlighted with a red frame, the second largest with an orange frame, and the third with a purple frame. ~~Red circle indicate mode that deviates from~~ These highlighted modes exhibit the ~~general trend~~ strongest response to the inclusion of the extreme-event penalisation term, exhibiting higher amplitude values in showing how the *Extreme extreme* DMD ~~compared~~ reweights the contribution of certain spatiotemporal structures to better capture high-magnitude temperature anomalies. Overall, the *Normal DMD* differences emphasise which modes gain importance when the reconstruction focuses on extremes rather than average system behaviour.

The most important mode (spatial pattern) and its corresponding temporal dynamics are presented. This mode is identified based on the largest amplitude difference, as shown in Fig. 16.

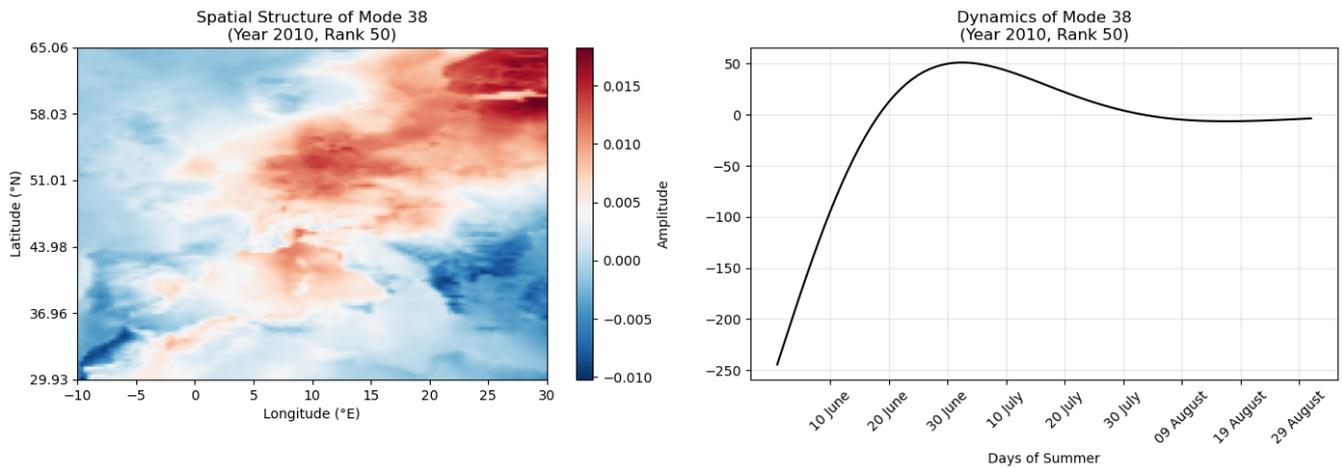


Figure 17. Spatial pattern and the dynamics of the *1st most important mode*.

375 The spatial pattern of this particular *1st most important mode* reveals the spatial pattern that strongly influences the average
 summer in Europe, but was less pronounced during the 2010 heatwave. It suggests a particularly strong and long-lasting
 blocking pattern developed over Russia. This blocking event was particularly intense from early July to mid-August 2010
 (Schaller et al., 2018), which is also confirmed by the dynamics showed on the right panel. The dynamics exhibited a marked
 increase at the beginning of summer (during June), followed by a relatively stable period with minimal oscillations throughout
 380 the remaining summer months.

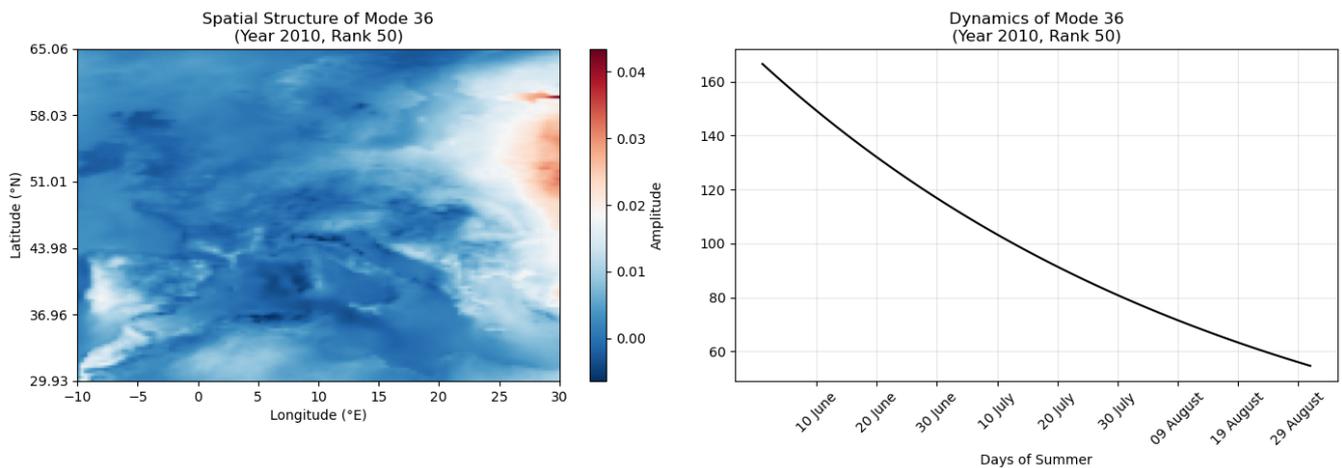


Figure 18. Spatial pattern and dynamics of the mode exhibiting the greatest increase in ~~Optimized-optimised~~ \mathbf{b} ($\mathbf{b}_{extreme}$) value relative to the Normal \mathbf{b} (\mathbf{b}_{normal}) value (red circle in the Fig. 16).

Fig. 18 shows the spatial pattern and temporal dynamics of the mode exhibiting the largest increase in ~~Optimized-optimised~~ \mathbf{b} ($\mathbf{b}_{extreme}$) compared to the ~~Normal-normal~~ \mathbf{b} (\mathbf{b}_{normal}). This mode captures a pronounced anomaly contrast between Eastern

Europe (extending into Western Asia) and the rest of the continent, suggesting a strong contribution to the 2010 heatwave. The associated dynamics reveal that this spatial pattern was most prominent at the beginning of the summer and gradually weakened as the season progressed.

4 Discussion

Until now, reduced-order techniques have primarily focused on capturing the average ~~behavior~~ behaviour of a system by filtering out noise and retaining only the dominant ~~,-“typical”-~~ modes. In contrast, we challenge this perspective by asking whether the method can be reversed to reveal the exceptional modes that characterize extreme events. Hence, we introduce a theoretical framework aimed at identifying outliers together with their associated spatiotemporal modes. This allows for a deeper understanding of the underlying dynamics and helps to uncover the drivers of extreme events within the climate system. The results presented here demonstrate that the proposed framework more effectively identifies outliers compared to the standard DMD method.

The main contribution of this work is a theoretical framework specifically designed to better detect outliers and approximate extreme - rather than average - ~~behavior~~ behaviour.

Our experiments clearly show that the original signal contains significantly more noise than the reconstructed fields, particularly as we move further forward in time. This is consistent with strong performance of the Koopman operator on a local scale. However, the primary objective of these experiments was to assess the effectiveness of the proposed DMD variation in accurately capturing extreme anomalies.

Given the current excitement surrounding large language models (LLMs), it is worth to mention that LLM architecture is equivalent to the Koopman operator-based architecture. However Koopman Operator Theory offers a robust framework for unsupervised learning using small amount of data, facilitating self-supervised learning of generative models that aligns more closely with human learning theories compared to some machine learning approaches (Mezić, 2023). Furthermore, by looking at significant modes (as in Fig. 10, Fig. 11, etc.) a physical understanding of driving patterns can be derived from DMD.

To validate the robustness of our event definition, we conducted a sensitivity analysis using 90th, 95th, and 99th percentile thresholds for extreme temperature anomalies. The resulting spatial modes and their temporal evolution remained consistent across all thresholds, indicating that the *extreme* DMD framework is largely insensitive to the specific percentile used. This robustness supports the rationality of the selection criterion and confirms that the essential spatiotemporal dynamics of extremes are well captured within the identified modal space, even when events outside this subspace are considered.

One of the central concern of every reduced order technique is the decision of the number of modes needed to represent the original signal. It is always a trade off between the accuracy and the complexity. Naturally, the higher the number of modes used in the reconstruction, the better the reconstruction will be (in terms of lower mean square error). However, increasing the complexity - represented by the rank parameter in our DMD algorithm - risks introducing noise into the reconstruction while also significantly raising computational cost. Our method was tested across various ranks, with all experiments consistently showing improvements when the ~~penalization~~ penalisation term was added. So, we can safely claim that the method is robust

with respect to the selection of the size of the rank. Unlike most of the other statistical approaches, we use rank as the only parameter, which allows for better interpretation of the model and more effective extraction of spatiotemporal patterns related to climate extremes. [Sensitivity analysis of rank choice is provided in the Supplementary Material.](#)

420 ~~Given the current excitement surrounding large language models (LLMs), it is worth to mention that LLM architecture is equivalent to the Koopman operator-based architecture. However Koopman Operator Theory offers a robust framework for unsupervised learning using small amount of data, facilitating self-supervised learning of generative models that aligns more closely with human learning theories compared to some machine learning approaches (Mezić, 2023). Furthermore, by looking at significant modes (as in Fig. 10, Fig. 11, etc.) a physical understanding of driving patterns can be derived from DMD.~~

425 The proposed framework is designed to be applicable to a broad range of climate extremes. While the current experiments have focused on temperature anomalies, specifically heatwave events, the approach can be readily extended to other types of extremes such as cold spells, heavy precipitation.

430 [Although the present study focuses on reconstructing extreme events, the *extreme* DMD framework could be extended to short-term forecasting. Once the spatiotemporal modes and amplitudes are identified, the system's future state can be projected by evolving the amplitudes according to the DMD linear dynamics. This approach could provide early-warning indicators by highlighting modes associated with high-amplitude extremes.](#)

435 [Practical challenges must be considered. The linear evolution assumption may limit predictive skill for highly nonlinear events, and mode selection or amplitude initialization requires careful calibration when applying the model to new conditions or different climate regimes. Real-time forecasting would also require near-real-time data assimilation. Nevertheless, integrating *extreme* DMD with ensemble methods or existing forecasting frameworks has the potential to enhance the prediction of extreme climate events, supporting risk assessment and mitigation strategies.](#)

440 [The present study primarily focuses on the methodological extension of DMD for capturing extreme events. Although the extracted modes are dynamically coherent, a detailed physical interpretation—such as linking these modes to specific circulation anomalies, blocking events, or teleconnection patterns—remains an open question. Previous studies \(e.g., Froyland et al. \(2021\), Chowdary et al. \(2023\), Zhang et al. \(2025\)\) have demonstrated that DMD and related Koopman-based approaches can reveal physically interpretable climate structures, suggesting that similar connections may be established here. Investigating such relationships will be an important avenue for future work aimed at bridging data-driven decomposition with atmospheric dynamics.](#)

5 Conclusion

445 This work introduces a novel Dynamic Mode Decomposition (DMD) variation [designed-tailored](#) to improve the reconstruction of extreme events [while providing a method-and](#) to extract spatiotemporal patterns (modes) specifically relevant to such extremes. ~~Using data-driven Koopman analysis, we demonstrate clear patterns of extremes and identify coherent behaviors in complex systems through spectral analysis of the Koopman operator. Assuming the system is well-behaved, with a diagonalizable Koopman operator and sufficient basis functions for reconstruction, we obtain a one-to-one correspondence between system~~

dynamics and observable evolution. This approach enables us to extract interpretable coherent spatiotemporal patterns. Our study makes three key contributions: (a) Regularisation design: We incorporate an extreme-event penalisation term into the DMD objective function, allowing the reconstruction to emphasize rare but important events. (b) Mode selection: The framework enables automated identification of modes most relevant to extremes, providing a systematic approach to isolate the spatial patterns driving extreme behaviour. (c) Climate science insight: Applying this method to European heatwaves reveals coherent spatiotemporal structures and the distinct impacts of anthropogenic emissions on extreme events, improving understanding of climate variability. By relying solely on time-delayed observations, this data-driven method offers insights into the system's key dynamics, serving as both a diagnostic tool and a potential prognostic model for understanding and predicting system behavior.

We conclude that by inverting the usual focus of the method—from modeling the system's average behavior to emphasizing its exceptional dynamics—and by carefully refining the algorithm accordingly, extreme events can be modeled more accurately across several common metrics. Moreover, this approach By shifting the focus from modelling average system behaviour to emphasising exceptional dynamics, the method achieves more accurate reconstruction of extremes and enables the extraction of spatial patterns that contribute to the reconstruction of extremes, along with insights into interpretable spatial patterns and their temporal evolution.

Even though in this work we focus on the diagnostic application, the method could be easily used for future state prediction. This extension is planned for the Heatwaves were chosen as a representative example to demonstrate the methodology; however, the approach is not limited to heat extremes. Extending the framework to other types of climate extremes - such as cold spells or heavy precipitation - remains an important direction for future work. Diagnosing the evolution in time of extreme situations, one could detect reoccurring patterns and therefore predict upcoming extremes. Although our analysis focuses on diagnostic applications, the framework could also be adapted for short-term forecasting, offering a potential tool for anticipating extreme events based on emerging spatiotemporal patterns.

Code availability. The codes for generating the results are made by means of scripting Python software. All codes used in this study can be obtained from the corresponding author upon reasonable request.

Data availability. The ERA5 reanalysis data used in this study are publicly available from the Copernicus Climate Data Store at <https://cds.climate.copernicus.eu/datasets>.

475 **Appendix A: Sensitivity to Rank Selection**

To assess the robustness of the proposed method with respect to the selected truncation rank (SVD Rank), we evaluated reconstruction performance across a range of ranks commonly used in DMD analyses. Figure A1 shows the mean squared

480 reconstruction error at extreme spatiotemporal points (MSE (extremes)) for both the standard DMD and the *extreme* DMD. Across all tested rank values, *extreme* DMD consistently achieved lower error than standard DMD. This demonstrates that the improved representation of extreme events is not dependent on a particular choice of rank and holds over a broad range of truncations. In addition, the plot shows that increasing the rank leads to systematically lower MSE values for both methods, with the gap between the two approaches widening slightly at higher ranks. This indicates that our method benefits from richer modal information and that the identification of extreme - relevant modes and the resulting reconstruction improvements are stable—and even strengthened - when the model order increases.

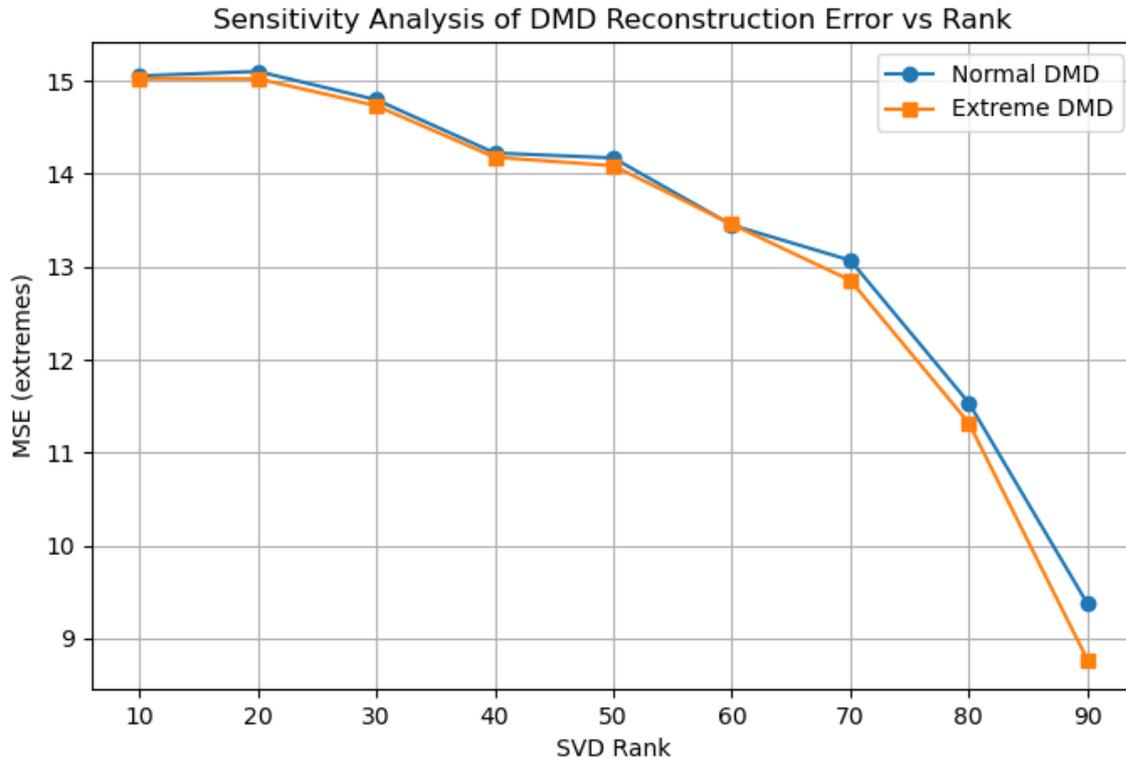


Figure A1. Sensitivity of extreme-event reconstruction to truncation rank. Across all tested SVD ranks, *extreme* DMD yields lower (MSE extremes) than standard DMD, indicating robustness to model order selection.

485 *Author contributions.* MA conducted the research, performed the analyses, and wrote the initial draft of the manuscript. JB provided close supervision throughout the project, contributed to the conceptual development, and substantially revised the manuscript. JS provided additional guidance and contributed to the manuscript revision.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. Authors acknowledge funding and support from the University of Hamburg. [JS and JB acknowledge funding from the Deutsche Forschungsgemeinschaft \(DFG, German Research Foundation\) under Germany's Excellence Strategy—EXC 2037: "CLICCS-Climate, Climatic Change, and Society" - Project number: 390683824, contribution to the Center for Earth System Research and Sustainability \(CEN\) at the University of Hamburg, Germany.](#) MA is grateful to the NumGeo and the Climate Extremes research group for valuable input and discussions throughout the development of this work. Additional thanks go to Dallas Murphy for his insightful advice on scientific writing, and to the SICCS Graduate School for their support and guidance. The authors also thank the Copernicus Climate Change Service for providing the ERA5 reanalysis data. Some language editing and writing suggestions were assisted by ChatGPT (OpenAI).

References

- Alcayaga, L., Larsen, G. C., Kelly, M., and Mann, J.: Identification of large-scale atmospheric structures under different stability conditions using Dynamic Mode Decomposition, *Journal of Physics: Conference Series*, 2265, 022 006, <https://doi.org/10.1088/1742-6596/2265/2/022006>, 2022.
- 500 Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe, *Science*, 332, 220–224, <https://doi.org/10.1126/science.1201224>, 2011.
- Berkooz, G., Holmes, P., and Lumley, J. L.: The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows, *Annual Review of Fluid Mechanics*, 25, 539–575, <https://doi.org/10.1146/annurev.fl.25.010193.002543>, 1993.
- Brunton, S. L., Proctor, J. L., and Kutz, J. N.: Discovering governing equations from data by sparse identification of nonlinear dynamical
505 systems, *Proceedings of the National Academy of Sciences*, 113, 3932–3937, <https://doi.org/10.1073/pnas.1517384113>, 2016.
- Brunton, S. L., Budišić, M., Kaiser, E., and Kutz, J. N.: Modern Koopman Theory for Dynamical Systems, <https://arxiv.org/abs/2102.12086>, arXiv:2102.12086 [math.DS], 2021.
- Casati, B., Yagouti, A., and Chaumont, D.: Regional Climate Projections of Extreme Heat Events in Nine Pilot Canadian Communities for
510 Public Health Planning, *Journal of Applied Meteorology and Climatology*, 52, 2669 – 2698, <https://doi.org/10.1175/JAMC-D-12-0341.1>, 2013.
- Chowdary, P. N., P. S., U. P., S. R., V. S., A. G. E., and M. D.: Unleashing the Power of Dynamic Mode Decomposition and Deep Learning for Rainfall Prediction in North-East India, <https://arxiv.org/abs/2309.09336>, arXiv:2309.09336 [cs.LG], 2023.
- Duke, D., Soria, J., and Honnery, D.: An error analysis of the dynamic mode decomposition, *Experiments in Fluids*, 52, 529–542, 2012.
- Erichson, N. B., Brunton, S. L., and Kutz, J. N.: Compressed dynamic mode decomposition for background modeling, *Journal of Real-Time
515 Image Processing*, 16, 1479–1492, <https://doi.org/10.1007/s11554-016-0655-2>, 2016.
- Froyland, G., Giannakis, D., Lintner, B. R., Pike, M., and Slawinska, J.: Spectral analysis of climate dynamics with operator-theoretic approaches, *Nature Communications*, 12, <https://doi.org/10.1038/s41467-021-26357-x>, 2021.
- García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., and and, E. M. F.: A Review of the European Summer Heat Wave of 2003, *Critical
Reviews in Environmental Science and Technology*, 40, 267–306, <https://doi.org/10.1080/10643380802238137>, 2010.
- 520 Gaspard, P.: *Chaos, Scattering and Statistical Mechanics*, Cambridge University Press, ISBN 978-0-511-62885-6, <https://doi.org/10.1017/CBO9780511628856>, 1998.
- H. Tu, J., W. Rowley, C., M. Luchtenburg, D., L. Brunton, S., and Nathan Kutz, J.: On dynamic mode decomposition: Theory and applications, *Journal of Computational Dynamics*, 1, 391–421, <https://doi.org/10.3934/jcd.2014.1.391>, 2014.
- Hersbach, H. and Dee, D.: ERA5 reanalysis is in production, *ECMWF Newsletter*, Vol. 147, Reading, United Kingdom, 370, 2016.
- 525 Holmes, P.: *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge university press, 2012.
- Hotelling, H.: Analysis of a Complex of Statistical Variables into Principal Components, *Journal of Educational Psychology*, 24, 417–441, <https://doi.org/10.1037/h0071325>, 1933.
- Jovanović, M. R., Schmid, P. J., and Nichols, J. W.: Sparsity-promoting dynamic mode decomposition, *Physics of Fluids*, 26, <https://doi.org/10.1063/1.4863670>, 2014.
- 530 Koopman, B. O.: Hamiltonian Systems and Transformation in Hilbert Space, *Proceedings of the National Academy of Sciences*, 17, 315–318, <https://doi.org/10.1073/pnas.17.5.315>, 1931.

- Lorenz, E.: Empirical Orthogonal Functions and Statistical Weather Prediction, Scientific report, Massachusetts Institute of Technology, Department of Meteorology, 1956.
- Lucarini, V., Faranda, D., Freitas, A. C. M., Freitas, J. M., Kuna, T., Holland, M., Nicol, M., Todd, M., and Vaienti, S.: Extremes and
535 Recurrence in Dynamical Systems, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, Wiley, ISBN 9781118632192, <https://books.google.hr/books?id=ebTOoQEACAAJ>, 2016.
- Mankovich, N., Bouabid, S., Nowack, P., Bassotto, D., and Camps-Valls, G.: Analyzing climate scenarios using dynamic mode decomposition with control, *Environmental Data Science*, 4, e16, <https://doi.org/10.1017/eds.2025.8>, 2025.
- Mezić, I.: Spectral Properties of Dynamical Systems, Model Reduction and Decompositions, *Nonlinear Dynamics*, 41, 309–325,
540 <https://doi.org/10.1007/s11071-005-2824-x>, 2005.
- Mezić, I.: Analysis of fluid flows via spectral properties of the Koopman operator, *Annu. Rev. Fluid Mech.*, 45, 357–378, 2013.
- Mezić, I.: Operator is the Model, <https://arxiv.org/abs/2310.18516>, arXiv:2310.18516 [math.DS], 2023.
- Navarra, A., Tribbia, J., Klus, S., and Lorenzo-Sánchez, P.: Variability of SST through Koopman Modes, *Journal of Climate*, 37, 4095 – 4114, <https://doi.org/10.1175/JCLI-D-23-0335.1>, 2024.
- 545 Perkins-Kirkpatrick, S. E. and Gibson, P. B.: Changes in regional heatwave characteristics as a function of increasing global temperature, *Scientific Reports*, 7, 12 256, <https://doi.org/10.1038/s41598-017-12520-2>, 2017.
- Proctor, J. L. and Eckhoff, P. A.: Discovering dynamic patterns from infectious disease data using dynamic mode decomposition, *International Health*, 7, 139–145, <https://doi.org/10.1093/inthealth/ihv009>, _eprint: <https://academic.oup.com/inthealth/article-pdf/7/2/139/16837591/ihv009.pdf>, 2015.
- 550 Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., and Henningson, D. S.: Spectral analysis of nonlinear flows, *J. Fluid Mech.*, 641, 85–113, 2009a.
- Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., and Henningson, D. S.: Spectral analysis of nonlinear flows, *Journal of Fluid Mechanics*, 641, 115–127, <https://doi.org/10.1017/S0022112009992059>, edition: 2009/11/18 Publisher: Cambridge University Press, 2009b.
- Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N.: Data-driven discovery of partial differential equations, <https://arxiv.org/abs/1609.06401>, arXiv:1609.06401 [nlin.PS], 2016.
- 555 Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., and Russo, S.: Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles, *Environmental Research Letters*, 13, 054 015, <https://doi.org/10.1088/1748-9326/aaba55>, 2018.
- Schmid, P. J.: Dynamic mode decomposition of numerical and experimental data, *Journal of fluid mechanics*, 656, 5–28, 2010.
- 560 Schmid, P. J. and Sesterhenn, J.: On dynamic mode decomposition: Theory and applications, 61st Annual Meeting of the APS Division of Fluid Dynamics, 53, 391–421, 2008.
- Schmid, P. J., Violato, D., and Scarano, F.: Decomposition of time-resolved tomographic PIV, *Experiments in fluids*, 52, 1567–1579, 2012.
- Shi, L., Haseli, M., Mamakoukas, G., Bruder, D., Abraham, I., Murphey, T., Cortes, J., and Karydis, K.: Koopman Operators in Robot Learning, arXiv preprint arXiv:2408.04200, <https://arxiv.org/abs/2408.04200>, accessed: 2025-01-24, 2024.
- 565 Smith, T. R., Moehlis, J., and Holmes, P.: Low-Dimensional Modelling of Turbulence Using the Proper Orthogonal Decomposition: A Tutorial, *Nonlinear Dynamics*, 41, 275–307, <https://doi.org/10.1007/s11071-005-2823-y>, 2005.
- Uchida, T., Yadidya, B., Lapo, K. E., Xu, X., Early, J. J., Arbic, B. K., Menemenlis, D., Hiron, L., Chassignet, E. P., Shriver, J. F., and Buijisman, M. C.: Dynamic Mode Decomposition of Geostrophically Balanced Motions From SWOT Cal/Val in the Separated Gulf Stream,

- Earth and Space Science, 12, e2024EA004 079, <https://doi.org/https://doi.org/10.1029/2024EA004079>, e2024EA004079 2024EA004079, 570 2025.
- Vogel, M. M., Zscheischler, J., Fischer, E. M., and Seneviratne, S. I.: Development of Future Heatwaves for Different Hazard Thresholds, *Journal of Geophysical Research: Atmospheres*, 125, e2019JD032 070, <https://doi.org/https://doi.org/10.1029/2019JD032070>, e2019JD032070 10.1029/2019JD032070, 2020.
- Wang, H., Liang, W., Liang, B., Ren, H., Du, Z., and Wu, Y.: Robust Position Control of a Continuum Manipulator Based on Selective 575 Approach and Koopman Operator, *IEEE Transactions on Industrial Electronics*, PP, 1–10, <https://doi.org/10.1109/TIE.2023.3236082>, 2023.
- Xu, Z., FitzGerald, G., Guo, Y., Jalaludin, B., and Tong, S.: Impact of heatwave on mortality under different heatwave definitions: A systematic review and meta-analysis, *Environment International*, 89-90, 193–203, <https://doi.org/https://doi.org/10.1016/j.envint.2016.02.007>, 2016.
- 580 Yin, C., Yang, Y., Chen, X., Yue, X., Liu, Y., and Xin, Y.: Changes in global heat waves and its socioeconomic exposure in a warmer future, *Climate Risk Management*, 38, 100 459, <https://doi.org/https://doi.org/10.1016/j.crm.2022.100459>, 2022.
- Zhang, Q., Liu, Y., and Wang, S.: The identification of coherent structures using proper orthogonal decomposition and dynamic mode decomposition, *Journal of Fluids and Structures*, 49, 53–72, <https://doi.org/https://doi.org/10.1016/j.jfluidstructs.2014.04.002>, 2014.
- Zhang, Z., Susuki, Y., and Okazaki, A.: Extracting transient Koopman modes from short-term weather simulations with sparsity-promoting 585 dynamic mode decomposition, <https://arxiv.org/abs/2506.14083>, 2025.