

基于动态模式分解的时间序列的 因果关系分析

汪荟宇

上海理工大学管理学院, 上海

收稿日期: 2024年1月22日; 录用日期: 2024年3月20日; 发布日期: 2024年3月27日

摘要

本文基于动态模式分解(DMD)分析四种气候时间序列(地表温度, 海面温度, 太阳黑子和碳排放)的模式之间的因果关系。通过延迟坐标的形式把每条单变量时间序列嵌入至多维, DMD捕捉了气候时间序列中不同的演化特征, 对这些特征进行不同数量的截断, 提取出不同快慢层面的模式。转移熵(TE)得出了模式之间的因果关系, 我们观察到因果关系在各种模式间的分布。总体上看, 由最大特征值提取的主导模式具有较大周期或长期趋势, 它作为很强的驱动因素, 与自身内部的快模式, 以及其他气候对象的模式之间有着很强的关联。太阳黑子的长期趋势对海面温度的各种周期和长期趋势有很大贡献, 影响程度大于地表温度。碳排放主要影响气温极慢的模式即超长期趋势, 气温反过来对碳排放的影响偏弱。四个气候对象构建的因果框架显示, 因果信息的传递基本平衡, 海面温度比地表温度吸收了更多的信息。这些方法的联合使用可以成为从复杂系统中提取各种模式, 分析内蕴性质和探索因果关系的灵活工具。

关键词

动态模式分解, 模式, 转移熵, 因果关系

Causal Relationship of Time Series Based on Dynamic Mode Decomposition

Huiyu Wang

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Jan. 22nd, 2024; accepted: Mar. 20th, 2024; published: Mar. 27th, 2024

Abstract

This study explores the causal relationships between modes in four climate time series (surface

文章引用: 汪荟宇. 基于动态模式分解的时间序列的因果关系分析[J]. 理论数学, 2024, 14(3): 58-73.

DOI: 10.12677/pm.2024.143085

temperature, sea surface temperature, sunspots and carbon emissions) using Dynamic Mode Decomposition (DMD). By embedding each univariate time series into multidimensional space using delay coordinates, DMD captures distinct evolutionary features in the climate time series. Different truncations of these features reveal modes at various timescales. Transfer Entropy (TE) is employed to quantify the causal relationships between modes, revealing the distribution of causal interactions among different modes. Overall, dominant modes extracted from the maximum eigenvalue exhibit larger cycles or long-term trends. These dominant modes serve as strong driving factors, displaying significant associations with fast internal modes and modes of other climate objects. The long-term trend in sunspots contributes significantly to various cycles and long-term trends in sea surface temperature, with a greater impact than on surface temperature. Carbon emissions predominantly influence extremely slow temperature modes, representing ultra-long-term trends, while the influence of temperature on carbon emissions is relatively weak. The causal framework constructed for the four climate objects illustrates a relatively balanced transfer of causal information, with sea surface temperature absorbing more information than surface temperature. The combined use of these methods serves as a flexible tool for extracting diverse modes from complex systems, analyzing intrinsic properties, and exploring causal relationships.

Keywords

DMD, Mode, TE, Causal Relationship

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

时间序列是复杂系统演化的记录，它蕴藏着系统内在的动力学特征。时间序列分析可以获取有意义的信息，比如捕捉系统的演化行为，提取系统中不同时空结构以及分析它们的相互作用。我们将时间序列视为一种信号，这个信号处理问题通常将数据分解为不同层面，从而将信号中的趋势与噪声解耦。过去的研究在建模和分析方面遇到了很多困难和挑战，高维和非线性特征使之在空间和时间上都表现出多尺度现象。尽管时间序列提供了丰富的可用数据，但构建准确的模型仍然是一项艰巨的工作。为了应对这种复杂性，本文提出了一种纯粹数据驱动的方法：动态模式分解(dynamic mode decomposition, DMD)。

DMD 起源于计算流体力学(CFD)领域，最初被设计为一种基于时空相干结构将复杂流体分解为低秩特征方法[1]。如今它被应用于广泛的领域，在计算机视觉中，DMD 能够将视频稳健地分离为背景和前景[2]。在生物识别领域，它有助于检测样本的有效性[3]。在机器人技术和神经科学领域，DMD 在估计人机交互中的扰动[4]和从大规模神经记录中提取一致模式[5]等方面发挥了关键作用。所有上述 DMD 应用都以数据矩阵作为输入，例如用 DMD 分析视频时，将输入数据转换为列矢量，每个列矢量表示一帧图像。

除了数据驱动性，DMD 还有一些其他的特点。首先是优秀的降维效果。随着一项 DMD 应用于单变量时间序列的研究问世[6]，体现了 DMD 能够从单变量序列中提取内蕴模式的能力。本文将用 DMD 处理不同的单变量信号，以延迟坐标的形式把单变量时间序列嵌入至多维时间序列，我们通过类比 DMD 如何处理视频来分析 DMD 如何处理这个数据矩阵：每一列表示当前时间点包含的观测量(像素)，每一行

代表一个时间窗口(帧数), 行数远大于列数。DMD 作为降维算法, 通过适当地投影, 它从大量的观测量中提取到少数特征构成的模式。第二, DMD 能够捕捉信号中主导的模式。模式的快慢程度主要受特征值控制, 其中上升和下降趋势由特征值的实部体现, 而周期性由特征值的虚部反映。DMD 允许我们自由控制特征截断数 r 来得到不同快慢层面的特征, 这样我们能够提取出一系列主要的模式, 并进一步发掘隐藏在部分模式中的非线性关联。我们不对原始时间序列作任何预处理, 一方面是因为 DMD 的数据驱动和无方程性质, 不对源数据做任何适用条件和限制; 另一方面是在剔除噪声的同时可能会导致原有的谐波结构丢失。与其他的滤波工具类似, DMD 保持了信号中的各种谐波结构, 这些结构作为模式的一部分, 它们是不可分离的, 共同构成了系统的一部分演化特征。本文的 DMD 框架把噪声视为不希望保留的分量, 通过选取不同的 r , 我们就过滤了不同层级的噪声, 提取了各种层面的模式。第三, DMD 作为一种线性方法, 它假设系统中存在一个不变的线性算子, 详见方法中对 DMD 的数学推导部分。

我们的研究对象是实际的气候时间序列。气候作为典型的复杂系统, 它受到各种不同环境因素的影响, 可大致分为两个领域: 自然效应和人类活动。碳排放已经成为人类活动的代表性驱动因素, 它对地球气候产生了重大影响; 然而气候与太阳的关系一直以来争论颇多, 作为太阳活动最直接的体现, 太阳黑子数量和气温的联系是过去研究的热点[7] [8]。越来越多的证据表明太阳黑子和气温之间的关系具有非线性特征[9] [10], 仅仅从线性的角度和方法难以捕捉到隐藏在各种模式之间的非线性关联。且过去的研究针对不同气候对象的两两关系, 没有考虑到同一气候对象的内部模式之间的关联。事实上, 内部的各种模式并不是相互独立的, 通过研究内部模式的关联, 我们能够更全面地理解某种气候对象自身的演化规律。不同模式或频段之间的性质差异十分显著, 我们的研究将提取不同快慢程度的模式, 更深入地揭示这些关联如何分布在各种模式中。

信息论方法在解释气候变量关联方面展现出很大的潜力[11], 已有一项基于转移熵的研究分析了包括太阳和碳排放在内的等很多环境因素对气温的影响, 但这项研究直接对信号本身计算转移熵, 未分解或提取任何时空结构[12]。转移熵(transfer entropy, TE)作为非线性因果关系的度量, 它的优势在于因果关系的单向性。比如地球气温绝不会对太阳黑子有任何影响, 即使气温和碳排放之间存在双向关系, 因果强度也会存在差异, TE 能提供更准确的结果。我们借助 k 近邻(kNN)估计的 TE 量化各种模式之间的关联[13] [14]。

我们的目标是分析因果关系如何分布在快慢模式中。技术方案包括三个步骤。第一, 对本文的四串单变量时间序列分别用延迟坐标方法嵌入至多维时间序列。第二, 分别确定最大特征截断数, 用 DMD 提取一系列模式。第三, 用 TE 分析各种模式之间的非线性关联。观察因果强度在不同模式之间的分布, 评估因果关系发生的快慢层面。几个比较重要的结果是: 关联集中于少数特征提取的慢模式; 最大的主导模式作为因, 它会对包含更多周期的其他主要模式产生重大影响; 太阳黑子对气温在各种快慢层面上均有显著影响; 碳排放主要影响气温的长期趋势; 四个气候对象构建的因果框架显示, 信息的流动基本平衡。

2. 方法

2.1. 数据来源

我们的数据包括地表温度(Global land surface temperature, *LST*), 海面温度(sea surface temperature, *SST*), 太阳黑子数量(*sunspot*)和碳排放(*carbon*)。 *LST* 和 *SST* 分别来自于 Berkeley Earth [15]和 Met Office Hadley Centre [16]。这里的温度指温度异常值(temperature anomalies), 它表示偏离平均温度的值, 平均温度来自于 1961~1990。Met Office Hadley Centre 不建议使用实际温度, 因为温度记录是以经纬度划分的网

格形式的，未覆盖整个地表比如高纬度地区；而温度异常值是对所有网格中异常值的平均，不确定性和系统性的误差都比实际温度小得多。太阳黑子数量来自于 SILSO [17]。碳排放来自于 Global Carbon Budget 2022，包括化石燃料的燃烧和氧化，水泥和船用燃料的排放，单位为百万吨[18]。除了碳排放为年度时间序列，其他都为月度时间序列。当我们分析年度和月度序列之间关联时，把月度序列平均化以匹配年度序列。所有序列的时间跨度都为 1850~2022，共 173 年。

2.2. 嵌入至多维时间序列

我们从四个单变量时间序列中任取一个长度为 n 的信号 \mathbf{x} ,

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T, \mathbf{x} \in \{LST, SST, sunspot, carbon\}$$

由相空间重构理论，以延迟坐标的形式把相邻样本嵌入到一个足够大的窗口。令窗口长度为 L ，可以得到一个长度为 $K = n - L + 1$ 的多维时间序列 \mathbf{X}_L 。对于三個月度时间序列，我们取 $L = 144$ ；对于年度的碳排放时间序列，取 $L = 3$ 。

$$\mathbf{X}_L = [\mathbf{x}_1^{lag}, \dots, \mathbf{x}_L^{lag}]$$

\mathbf{X}_L 中的列称为滞后向量 \mathbf{x}_t^{lag} 。

$$\mathbf{x}_t^{lag} = [x_t, x_{t+1}, \dots, x_{n-L+t}]^T, 1 \leq t \leq L$$

窗口的个数即 \mathbf{x}_t^{lag} 的长度为 $n - L + 1$ ，因此单变量时间序列 \mathbf{x} 被分解为长度为 L 的 $n - L + 1$ 个重叠窗口。水平堆叠 \mathbf{x}_t^{lag} ， \mathbf{X}_L 可展开为矩阵的形式。

$$\mathbf{X}_L = [\mathbf{x}_1^{lag}, \dots, \mathbf{x}_L^{lag}] = \begin{bmatrix} x_1 & x_2 & \dots & x_L \\ x_2 & x_3 & \dots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \dots & x_n \end{bmatrix}$$

2.3. 动态模式分解

\mathbf{x} 中快照的采样间隔是固定的，相邻的滞后向量 $(\mathbf{x}_t^{lag}, \mathbf{x}_{t+1}^{lag})$ 也有着固定的时间间隔 Δt 。假设存在一个不变的线性算子 \mathbf{A} 作用于 $(\mathbf{x}_t^{lag}, \mathbf{x}_{t+1}^{lag})$ 中。

$$\mathbf{x}_{t+1}^{lag} = \mathbf{A}\mathbf{x}_t^{lag}$$

\mathbf{A} 对任意的 $(\mathbf{x}_t^{lag}, \mathbf{x}_{t+1}^{lag})$ 都成立，我们可以把 \mathbf{A} 作用前后的 \mathbf{x}_t^{lag} 和 \mathbf{x}_{t+1}^{lag} 拓宽成两个滑动矩阵 \mathbf{X}_1 和 \mathbf{X}_2 。 \mathbf{A} 把 \mathbf{X}_1 映射到 \mathbf{X}_2 ，捕获了 \mathbf{X}_L 所有的 $(\mathbf{x}_t^{lag}, \mathbf{x}_{t+1}^{lag})$ 。

$$\mathbf{X}_1 = [\mathbf{x}_1^{lag}, \dots, \mathbf{x}_{L-1}^{lag}]$$

$$\mathbf{X}_2 = [\mathbf{x}_2^{lag}, \dots, \mathbf{x}_L^{lag}]$$

$$\mathbf{X}_2 = \mathbf{A}\mathbf{X}_1$$

我们可以通过 \mathbf{X}_1 的 MP 伪逆 \mathbf{X}_1^+ 来间接地计算出 \mathbf{A} 。

$$\mathbf{A} = \mathbf{X}_2\mathbf{X}_1^+$$

对 \mathbf{X}_1 进行降阶奇异值分解(SVD)得到左右奇异矩阵和奇异值 \mathbf{U} , \mathbf{V} , Σ ，这三个矩阵的正交性使得我

们可以直接得出 \mathbf{X}_1^+ 。

$$\mathbf{X}_1 \approx \mathbf{U}\Sigma\mathbf{V}^\dagger \Rightarrow \mathbf{X}_1^+ = \mathbf{V}\Sigma^{-1}\mathbf{U}^\dagger$$

完整大小的 \mathbf{A} 由 \mathbf{X}_2 和经过分解后的 \mathbf{X}_1^+ 得出

$$\mathbf{A} = \mathbf{X}_2\mathbf{V}\Sigma^{-1}\mathbf{U}^\dagger \in \mathbb{R}^{(n-L) \times (n-L)}$$

\mathbf{A} 由数个低秩矩阵得出，因此它的秩，或者说它代表的动力系统所蕴含的特征数量远低于它的实际尺寸。我们希望构造 \mathbf{A} 的低秩近似，这不但可以降低计算成本，更重要的是我们希望从 \mathbf{A} 中寻找低维底流形中的特征。 \mathbf{U} 作为原始数据堆叠而成的滑动矩阵 \mathbf{X}_1 的左奇异向量，它可以对 \mathbf{X}_1 的行数进行压缩：

$$\tilde{\mathbf{X}}_1 = \mathbf{U}^\dagger\mathbf{X}_1 \in \mathbb{R}^{L \times L}$$

\mathbf{U} 把 \mathbf{X}_1 中远大于列数的行数 ($n-L \gg L$) 压缩到和列数一样的水平上，降维效果十分可观。两个低维的滑动矩阵之间也应存在一个低维的不变算子，

$$\tilde{\mathbf{X}}_2 = \tilde{\mathbf{A}}\tilde{\mathbf{X}}_1 \in \mathbb{R}^{L \times L}$$

\mathbf{U} 对 \mathbf{X}_2 进行同样的投影，并利用 $\mathbf{A} = \mathbf{X}_2\mathbf{X}_1^+$ ， \mathbf{U} 将 \mathbf{A} 投影到低维 $L \times L$ 子空间，

$$\tilde{\mathbf{A}} \approx \mathbf{U}^\dagger\mathbf{A}\mathbf{U} = \mathbf{U}^\dagger\mathbf{X}_2\mathbf{V}\Sigma^{-1} \in \mathbb{R}^{L \times L}$$

通过保留前 r 个特征值与其对应的特征向量，我们就可以得到前 r 个最大的特征， \mathbf{A} 被投影到低维 $r \times r$ 子空间，我们现在可以改变 r 来调整特征的数量。

$$\mathbf{A}_r \approx \mathbf{U}_r^\dagger\mathbf{A}\mathbf{U}_r = \mathbf{U}_r^\dagger\mathbf{X}_2\mathbf{V}_r\Sigma_r^{-1} \in \mathbb{R}^{r \times r}$$

对 \mathbf{A}_r 特征分解 $\mathbf{A}_r\boldsymbol{\omega}_r = \boldsymbol{\Lambda}_r\boldsymbol{\omega}_r$ ，代入上式有

$$\mathbf{A}\mathbf{U}_r \approx \mathbf{U}_r\mathbf{A}_r = \mathbf{U}_r\boldsymbol{\omega}_r\boldsymbol{\Lambda}_r\boldsymbol{\omega}_r^{-1} \Rightarrow \mathbf{A}(\mathbf{U}_r\boldsymbol{\omega}_r) \approx (\mathbf{U}_r\boldsymbol{\omega}_r)\boldsymbol{\Lambda}_r$$

这样就得到了前 r 个特征 \mathbf{A}_r 重构的 \mathbf{A} 的特征向量 $\boldsymbol{\varphi}_r = \mathbf{U}_r\boldsymbol{\omega}_r$ ，我们在计算中应用它的对偶 [19] $\boldsymbol{\Phi}_r = \mathbf{A}\boldsymbol{\varphi}_r$ 作为特征向量。

$$\boldsymbol{\Phi}_r \triangleq \mathbf{X}_2\mathbf{V}_r\Sigma_r^{-1}\boldsymbol{\omega}_r$$

2.4. 提取模式

DMD 把离散时间动力系统 $\mathbf{x}_{t+1}^{lag} = \mathbf{A}\mathbf{x}_t^{lag}$ 中的不变算子 \mathbf{A} 推广至连续时间动力学 \mathbf{A} 并等间隔 Δt 采样，得到一个连续时间的动力系统

$$\frac{d\mathbf{x}^{lag}}{dt} = \mathbf{A}\mathbf{x}^{lag}, \mathbf{A} = \ln \mathbf{A} / \Delta t$$

我们用离散时间特征值 λ ，特征向量 $\boldsymbol{\Phi}$ 以及初始状态 \mathbf{x}_1^{lag} 对应的系数 $\mathbf{b} = \boldsymbol{\Phi}^+\mathbf{x}_1^{lag}$ 写出这个系统的离散形式的解

$$\tilde{\mathbf{x}}_t = \sum_{j=1}^L \boldsymbol{\Phi}_j \lambda_j^{t-1} b_j = \boldsymbol{\Phi}\boldsymbol{\Lambda}^{t-1}\mathbf{b}$$

它由 L 个特征解出，取前 r 个特征，将上式展开并用连续时间特征值 $\sigma = \ln \lambda / \Delta t$ 替代，得到由前 r 个特征值构成的向量 σ_r 对某条滞后向量 \mathbf{x}_t^{lag} 提取的 r 阶模式 \mathbf{x}_t^r ，如图 1：

$$\mathbf{x}_t^r = \sum_{k=1}^r \phi_k \exp(\sigma_k(t-1)) \mathbf{b}_k = \begin{matrix} \text{r列} \\ \left[\begin{array}{c} \phi_1 \\ \vdots \\ \phi_r \end{array} \right] \end{matrix} \begin{matrix} \left[\begin{array}{c} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_r \end{array} \right] \end{matrix} \stackrel{\circ}{=} \exp(\sigma_r(t-1)) = \begin{bmatrix} x_t^r \\ x_{t+1}^r \\ \vdots \\ x_{n-L+t}^r \end{bmatrix}$$

其中 $t = 1, \dots, L$, \circ 为 Hadamard 积

Figure 1. Mode extraction formula of a \mathbf{x}_t^{lag}

图 1. 对某条 \mathbf{x}_t^{lag} 的模式提取公式

\mathbf{x}_t^r 作为从 \mathbf{x}_t^{lag} 提取的 r 阶模式，它的元素和 r 有关，但长度不受特征截断数 r 的控制，仅仅取决于窗口长度 L 。每一条 \mathbf{x}_t^{lag} 都有其对应的模式，我们可以得到由所有 \mathbf{x}_t^{lag} 构成的矩阵形式的模式 \mathbf{X}_L^r 。其上下标 r 和 L 表示它是在一个特定的窗口长度下，由某个特征截断数生成的。

$$\mathbf{X}_L^r = [\mathbf{x}_1^r, \dots, \mathbf{x}_L^r]$$

注意到 \mathbf{X}_L 是 Hankel 矩阵，它的每一条反对角线都对应同一个快照，那么 \mathbf{X}_L^r 的每一条反对角线也应相等。事实上，滞后向量 \mathbf{x}_t^{lag} 张成了 Krylov 子空间：

$$\mathbf{X}_L = [\mathbf{x}_1^{lag}, \mathbf{A}\mathbf{x}_1^{lag}, \dots, \mathbf{A}^{L-1}\mathbf{x}_1^{lag}]$$

这等价于利用滑动矩阵进行一次滑动 $\mathbf{X}_2 = \mathbf{A}\mathbf{X}_1$ 。 \mathbf{X}_1 和 \mathbf{X}_2 捕捉了所有的 \mathbf{x}_t^{lag} ，保证了不变算子 \mathbf{A} 的准确性，但未保证 \mathbf{X}_L^r 是 Hankel 的。我们对 \mathbf{X}_L^r 中的每一条反对角线进行加权平均，[6] 得到最终的向量形式的模式 $\hat{\mathbf{x}}_L^r$ ，形象化的展示见图 2：

$$\hat{\mathbf{x}}_L^r = 1/\#(D_k) \sum_{i,j \in D_k} \mathbf{X}_L^r(i,j)$$

其中 i, j 为 \mathbf{X}_L^r 的行和列，

D_k 为第 k 条 \mathbf{X}_L^r 的反对角线，

$\#(D_k)$ 为 D_k 中元素的数量。

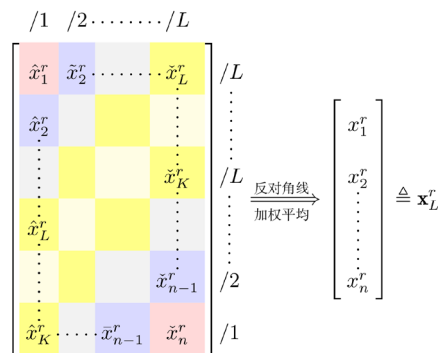


Figure 2. Anti-diagonal weighted average method, where different over lines indicate that samples come from different \mathbf{x}_t^{lag} , not necessarily equal

图 2. 反对角线加权平均法示意，不同的上划线表示这些样本来自不同的 \mathbf{x}_t^{lag} ，不一定相等

我们使用的 DMD 算法来自 Kutz *et al.* [20]主要步骤如表 1。

Table 1. Dynamic mode decomposition (DMD) algorithm
表 1. 动态模式分解(DMD)算法

算法：动态模式分解算法	
1. 计算 \mathbf{X}_1 的 SVD, 得线性算子 \mathbf{A} : $\mathbf{X}_1 = \mathbf{U}\Sigma\mathbf{V}^\dagger \Rightarrow \mathbf{X}_2 = \mathbf{A}\mathbf{U}\Sigma\mathbf{V}^\dagger \Rightarrow \mathbf{A} = \mathbf{X}_2\mathbf{V}\Sigma^{-1}\mathbf{U}^\dagger$	
2. 构造 \mathbf{A} 的前 r 个特征的低秩近似 \mathbf{A}_r : $\mathbf{A}_r \approx \mathbf{U}_r^\dagger\mathbf{A}\mathbf{U}_r = \mathbf{U}_r^\dagger\mathbf{X}_2\mathbf{V}_r\Sigma_r^{-1}$	
3. 对 \mathbf{A}_r 特征分解: $\mathbf{A}_r\boldsymbol{\omega}_r = \boldsymbol{\Lambda}_r\boldsymbol{\omega}_r$	
4. 构造 r 阶 DMD 特征向量 Φ_r 以及初始系数: $\Phi_r \triangleq \mathbf{X}_2\mathbf{V}_r\Sigma_r^{-1}\boldsymbol{\omega}_r$, $\mathbf{b}_r = \Phi_r^\dagger\mathbf{x}_1^{log}$	
5. 以前 r 个特征提取 \mathbf{x} 的模式并平均化: $\mathbf{x}_t^r = \sum_{k=1}^r \phi_k \exp(\sigma_k t) b_k \Rightarrow \mathbf{x}_L^r$	

我们希望从每种信号中提取出一系列模式, 由于特征值的分布不是均匀的, 所以需要遍历每一个 r 来提取模式, 共 R 个模式, 否则可能会遗漏合适的模式。确定模式的数量即最大的特征截断数 R 的方法是: 对 \mathbf{X}_1 进行 SVD, 当前 r 个奇异值的平方和达到 95% 的总平方和时, 取这个 r 作为最大的特征截断数 R 。表 2 为对各信号使用不同的最大特征截断数 R 来得到前 r 个特征值提取的模式。大于 R 的特征值, 即后 $L - r$ 个特征值表示对源信号的贡献小于 5% 的噪声, 通过在 r 处截断, 就剔除了这个 r 对应的贡献率以下的噪声, 提取出相比原信号更慢的模式。碳排放是年度信号, 其趋势非常平缓, 我们只需提取其一个模式。

Table 2. Extract different numbers of modes for each signal
表 2. 对各信号提取出不同数量的模式

提取模式: 特征截断数 r , 取前 r 个特征值提取模式	
LST	$r = 1:R$, $R = 1, 2, \dots, 96$, $L = 144$, 共 96 种模式
SST	$r = 1:R$, $R = 1, 2, \dots, 12$, $L = 144$, 共 12 种模式
太阳黑子	$r = 1:R$, $R = 1, 2, \dots, 12$, $L = 144$, 共 12 种模式
碳排放	$r = R = 1$, $L = 3$, 1 种模式

每个信号提取的 R 种不同的模式构成了一个模式集 \mathbf{X}^R , 其上标 R 区别于由特定的特征截断数 r 得到的 \mathbf{X}_L^r 。

$$\mathbf{X}^R = \{\mathbf{x}^1, \dots, \mathbf{x}^r, \dots, \mathbf{x}^R\}$$

2.5. 用转移熵计算模式间因果关系

对两个原始气候时间序列 U, V 依表 2 分别用 DMD 提取模式, 得 U^r 和 V^r 。用 TE 计算其中模式的关联, 得到一个由 TE 值组成的矩阵 $\mathbf{TE}_{U \rightarrow V}$ 。

$$\mathbf{TE}_{U \rightarrow V} = \{TE_{U \rightarrow V}^{m,n} \mid m \leq R(U), n \leq R(V)\}$$

其中 $TE_{U \rightarrow V}^{m,n} \triangleq TE(U^m \rightarrow V^n)$, $U^m \in U^{R(U)}$, $V^n \in V^{R(V)}$ 。

展开 $\mathbf{TE}_{U \rightarrow V}$, 随着特征截断数 r 的增加, 从左至右 V 的模式越来越快, 从上至下 U 的模式越来越快。图 3 为展开 $\mathbf{TE}_{U \rightarrow V}$ 的示例:

$$\mathbf{TE}_{U \rightarrow V} = \begin{matrix} U^1 \\ \vdots \\ U^m \\ \vdots \\ UR(U) \end{matrix} \begin{bmatrix} V^1 \dots V^n \dots V^{R(V)} \\ TE_{U \rightarrow V}^{1,1} \dots TE_{U \rightarrow V}^{1,n} \dots TE_{U \rightarrow V}^{1,R(V)} \\ \vdots \\ TE_{U \rightarrow V}^{m,1} \dots TE_{U \rightarrow V}^{m,n} \dots TE_{U \rightarrow V}^{m,R(V)} \\ \vdots \\ TE_{U \rightarrow V}^{R(U),1} \dots TE_{U \rightarrow V}^{R(U),n} \dots TE_{U \rightarrow V}^{R(U),R(V)} \end{bmatrix}$$

Figure 3. An example of constructing a transfer entropy matrix, each row or column is a specific mode, every element in the matrix is the transfer entropy value under this specific mode pair

图 3. 构造转移熵矩阵的示例，任意一行或一列都是一个特定的模式，矩阵中任意元素都是在这对特定模式下的转移熵值

对于可能存在的 12 种因果关系(包括三种月度气候信号内部模式之间的关系，太阳黑子不作为果，因为地球上的信号显然不可能对其造成影响)，我们可以构造一个 $\mathbf{TE}_{U \rightarrow V}$ 组成的集合 \mathbf{TE} 。按 LST , SST , 太阳黑子和碳排放的顺序，可以将 \mathbf{TE} 表示为一个 4×4 的分块矩阵：

$$\mathbf{TE} = \begin{bmatrix} \mathbf{TE}_{LST \rightarrow LST} & \mathbf{TE}_{SST \rightarrow LST} & \mathbf{TE}_{sunspot \rightarrow LST} & \mathbf{TE}_{carbon \rightarrow LST} \\ \mathbf{TE}_{LST \rightarrow SST} & \mathbf{TE}_{SST \rightarrow SST} & \mathbf{TE}_{sunspot \rightarrow SST} & \mathbf{TE}_{carbon \rightarrow SST} \\ / & / & \mathbf{TE}_{sunspot \rightarrow sunspot} & / \\ \mathbf{TE}_{LST \rightarrow carbon} & \mathbf{TE}_{SST \rightarrow carbon} & \mathbf{TE}_{sunspot \rightarrow carbon} & / \end{bmatrix}$$

每一对模式之间的转移熵值可以由 Shannon 熵 H 给出。

$$\begin{aligned} TE_{U \rightarrow V}^{m,n} &= -H(v_{t+1} | u_t, v_t) + H(v_{t+1} | v_t) \\ &= -H(v_{t+1}, u_t, v_t) + H(u_t, v_t) + H(v_{t+1}, v_t) - H(v_t) \\ &= \sum_{v_{t+1}, u_t, v_t} p(v_{t+1}, u_t, v_t) \log \frac{p(v_{t+1} | u_t, v_t)}{p(v_{t+1} | v_t)} \end{aligned}$$

其中 $u_t \in U^m$, $v_t \in V^n$ 。

我们以 v_t 为例阐释基于 kNN 的计算概率密度的方法。设 v_t 与其第 k 个最近的邻居之间的距离为 D_t 。 v_t 的邻域即以它作为中心， D_t 为直径的球中包含 s_t 个快照，则 v_t 在此处的概率密度可理解为

$$p(v_t) \sim \frac{s_t}{\mathcal{V}(\mathfrak{B}(v_t, D_t/2))}$$

$\mathcal{V}(\mathfrak{B}(v_t, D_t/2))$ 为 v_t 的球形邻域的体积。

它是这个区域中概率出现 s_t 个观测量的直观表示。具体地，设 $p_k(D_t)$ 为某快照出现在 v_t 邻域内的概率， p 为其概率密度，则

$$p_k(D_t) = \int_{\|\delta - v_t\| \leq D_t/2} p(\delta) d\delta$$

$\{v_t\}$ 共含 $N = n - 1$ 个快照(不包括 $t = n$ 时刻，同理 $\{v_{t+1}\}$ 不包括初始时刻)，除 v_t 自身外，剩余 $N - 1$

个快照中恰好有 k 个落在球内的概率可以由二项分布估计，

$$P_k(D_t) dD_t = \binom{N-1}{k} [p_k(D_t)]^k [1-p_k(D_t)]^{N-k-1} dp_k(D_t)$$

其中 $P_k(D_t)$ 为 v_t 与其第 k 个最近邻之间距离的分布，它是一个对 D_t 的估计。 $P_k(D_t) dD_t$ 等价于当球直径 D_t 增加一个很小的单位 dD_t 时，恰好把 k 个快照包含在 D_t 范围内的概率。两边同时乘以 $\log p_k(D_t)$ 并在 $(0, \infty)$ 上积分，显然左边的概率密度 $P_k(D_t)$ 归一了，留下 $\log p_k$ 的期望[21]:

$$\begin{aligned} \mathbb{E}(\log p_k) &= \int_0^\infty P_k(D_t) \log p_k(D_t) dD_t \\ &= \binom{N-1}{k} \int_0^1 p_k^k (1-p_k)^{N-k-1} \log p_k dp_k \\ &\approx \psi(k) - \psi(N) \end{aligned}$$

上式第三步的估计由 Beta 函数导出: $\partial_x B(x, y) = B(x, y)(\psi(x) - \psi(x+y))$ ，其中 ψ 为 digamma 函数，详见[22]。回顾概率的直观表示， p_k 即 v_t 的邻域内包含的快照个数 s_t ，kNN 假设它与邻域的大小 D_t 成正比，即 v_t 的概率密度 p 与 D_t 无关，我们就得到了 $H(v_t)$ 所需的概率密度:

$$p(v_t) = \frac{p_k(D_t)}{D_t^d}, \quad D_t^d \text{ 为无穷范数下的 } D_t \text{ 的体积}$$

对于我们的向量 $\{v_t\}$ ，维数 $d=1$ 。综合以上两式，得

$$\log p(v_t) \approx \psi(k) - \psi(N) - \mathbb{E}(\log D_t)$$

这样就得到了基于单变量最近邻距离 D_t 估计的 $H^1(v_t)$

$$\begin{aligned} H^1(v_t) &= -\int p(v_t) \log p(v_t) dt \\ &\approx -\frac{1}{N} \sum_{t=1}^N \log p(v_t) \\ &= \psi(N) - \psi(k) + \frac{1}{N} \sum_{t=1}^N \log D_t \end{aligned}$$

在实际的计算中，我们的最近邻距离 ϵ_t 的计算统一来自于三变量 (v_{t+1}, u_t, v_t) 构成的三维空间。令 $\mathbf{z}_t = [v_{t+1}, u_t, v_t]^T$ ，捕捉所有的 \mathbf{z}_t ，得到一个三维的时间序列 \mathbf{Z} 。

$$\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_N] = \begin{bmatrix} v_2 & \cdots & v_{N+1} \\ u_1 & \cdots & u_N \\ v_1 & \cdots & v_N \end{bmatrix}$$

寻找每个快照 \mathbf{z}_t 在 \mathbf{Z} 中最近的邻居 \mathbf{z}_t^* ，得到最近邻距离 ϵ_t

$$\epsilon_t = \|\mathbf{z}_t - \mathbf{z}_t^*\|_2, \quad t=1, \dots, N$$

接下来计算 $H(\mathbf{z}_t)$ 的方法与前文中计算 $H^1(v_t)$ 同理，它由联合概率密度导出。其中第三项分子上的 3 表示 \mathbf{z}_t 的维数 $d=3$

$$H(v_{t+1}, u_t, v_t) = \psi(N) - \psi(k) + \frac{3}{N} \sum_{t=1}^N \log \epsilon_t$$

为计算 $H(v_t)$, $H(u_t, v_t)$ 和 $H(v_{t+1}, v_t)$, 需要分别对向量 \mathbf{v} 和两个由 \mathbf{Z} 拆分而来的矩阵 \mathbf{Z}_{uv} , $\mathbf{Z}_{v'v}$ 进行分析,

$$\begin{aligned}\mathbf{v} &= [v_1, \dots, v_N] \\ \mathbf{Z}_{uv} &= \begin{bmatrix} u_1, \dots, u_N \\ v_1, \dots, v_N \end{bmatrix} \\ \mathbf{Z}_{v'v} &= \begin{bmatrix} v_2, \dots, v_{N+1} \\ v_1, \dots, v_N \end{bmatrix}\end{aligned}$$

设 v_t , (u_t, v_t) , (v_{t+1}, v_t) 在 ϵ_t 内分别含 s_t^v , s_t^{uv} , $s_t^{v'v}$ 个快照, 对任意 k , \mathbf{z}_t 构造的三维空间中的 ϵ_t 都会大于边缘空间 \mathbf{v} , \mathbf{Z}_{uv} , $\mathbf{Z}_{v'v}$ 中邻居的距离, s_t 相当于计算边缘概率密度所需的最近邻个数。

$$\begin{aligned}s_t^v &= \#\{v_j \mid \|v_j - v_t\| < \epsilon_t/2\} \\ s_t^{uv} &= \#\{(v_j, u_j) \mid \|(v_j, u_j) - (v_t, u_t)\| < \epsilon_t/2\} \\ s_t^{v'v} &= \#\{(v_{j+1}, v_j) \mid \|(v_{j+1}, v_j) - (v_{t+1}, v_t)\| < \epsilon_t/2\}\end{aligned}$$

其中 $\#\{\cdot\}$ 为元素的数量, $j \neq i$ 。

$H(v_t)$, $H(u_t, v_t)$ 和 $H(v_{t+1}, v_t)$ 如下, 由边缘概率密度导出。后者通过 $H^1(v_t)$ 中的二项分布方法, s_t^v , s_t^{uv} , $s_t^{v'v}$ 作为最近邻个数来估计。 $\log \epsilon_t$ 的平均需要乘以当前快照的维数 $d = 1$ 或 2 。由于 s_t^v , s_t^{uv} , $s_t^{v'v}$ 只搜索除自身之外满足条件的快照, 所以实际满足条件的快照数量需要各自加一。相比 $H(\mathbf{z}_t)$ 中的 $\psi(k)$, 它们用 $\psi(s_t + 1)$ 在所有 t 上的均值来代替, 这相当于在逐个快照中取 $s_t + 1$ 个被包含在 ϵ_t 球内的快照再平均化。

$$\begin{aligned}H(v_t) &= \psi(N) + \frac{1}{N} \sum_{t=1}^N \psi(s_t^v + 1) - \frac{1}{N} \sum_{t=1}^N \log \epsilon_t \\ H(u_t, v_t) &= \psi(N) + \frac{1}{N} \sum_{t=1}^N \psi(s_t^{uv} + 1) - \frac{2}{N} \sum_{t=1}^N \log \epsilon_t \\ H(v_{t+1}, v_t) &= \psi(N) + \frac{1}{N} \sum_{t=1}^N \psi(s_t^{v'v} + 1) - \frac{2}{N} \sum_{t=1}^N \log \epsilon_t\end{aligned}$$

这样我们就完成了转移熵 $TE_{U \rightarrow V}$ 的估计:

$$TE_{U \rightarrow V} = \psi(k) + \langle \psi(s_t^v + 1) - \psi(s_t^{uv} + 1) - \psi(s_t^{v'v} + 1) \rangle$$

本文取 $k = 1$ 计算 TE 值, 转移熵算法来自 Papanas [13]。表 3 为主要步骤。

Table 3. Transfer entropy (TE) algorithm

表 3. 转移熵(TE)算法

算法: 转移熵算法

1. 嵌入至 3 维时间序列 \mathbf{Z} , 计算最近邻距离 ϵ_t 。
2. 估计联合概率密度 $p(v_{t+1}, u_t, v_t)$ 。
3. 拆分 \mathbf{Z} , 用包含近邻的个数 s_t^v , s_t^{uv} , $s_t^{v'v}$ 估计边缘概率密度。
4. 计算概率密度对应的 Shannon 熵, 得到转移熵。

图 4 是本文的方法流程:

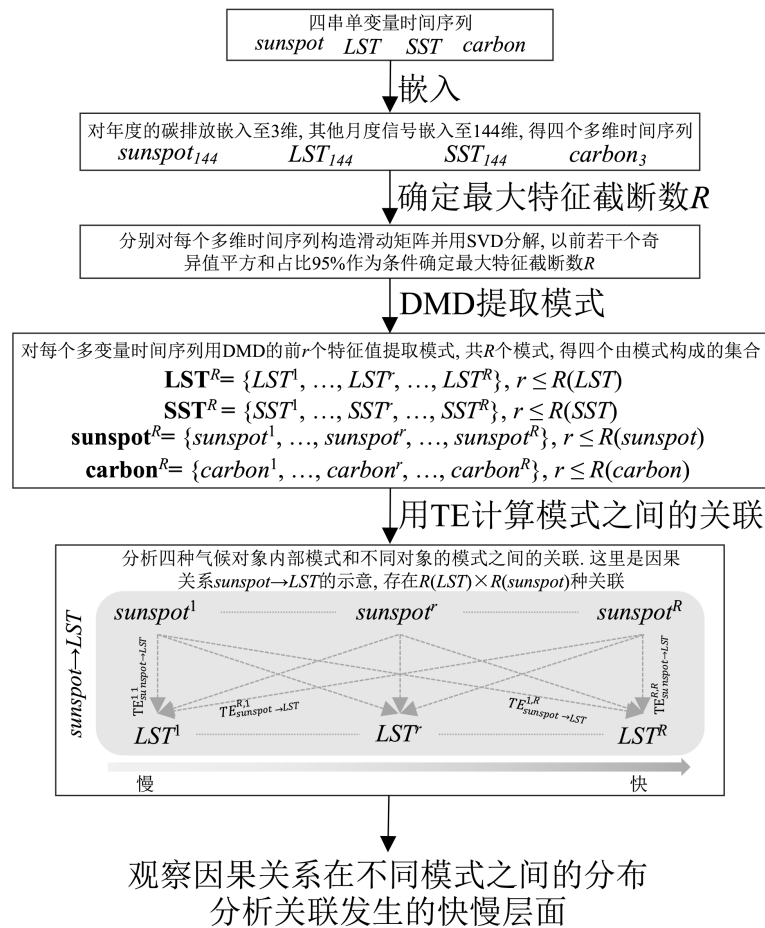


Figure 4. Flow chart for analyzing causality between climate signals and internal modes based on DMD

图 4. 基于 DMD 分析气候信号之间及内部模式关联强度的流程图

3. 结果

图 5 展示了三种月度信号分别嵌入至多维时间序列的奇异值与主要模式对应的特征值。第一列中的奇异值按降序排列, LST 的奇异值随个数的增加, 占有奇异值之比相较 SST 与太阳黑子更为缓慢。 LST 的最大奇异值贡献了 60%, 太阳黑子的最大奇异值贡献了 65%, 而 SST 的最大奇异值贡献了 85%, 这体现了 SST 相比其它两个月度信号噪声更小的特点。太阳黑子的最大奇异值占比与 LST 相似, 但前者与 SST 一样, 随着奇异值个数的增加, 前几个奇异值的贡献速度上升极快。 LST 的奇异值直到 48 个才达到 85%。第二列显示前几个奇异值的衰减比指数更快, 随后收敛。

第三列和第四列展示了 DMD 方法得到的线性算子 \mathbf{A} 的特征值, 它们反映了对应模式的时间演化。特征值虚部决定了模式的周期(快慢程度), 实部决定了模式的趋势(衰减速度)。因此, 较小虚部和正实部的特征值对应的模式体现了信号的宏观趋势; 大虚部对应的模式捕捉了信号的微观行为或噪声, 因为大虚部意味着快速振荡。显示了所有 144 个特征值(空心圆), 其中若干种主要模式对应的特征值用叉号标记。特征值基本分布在单位圆上, 有极少数几个例外, 相对于虚轴是对称的。特征值的实部都不大于 1, 即系统是稳定的; 且三种信号都存在一个特征值为 1 对应的模式, 这个模式是一种稳定的状态, 因其不包含任何衰减和振荡。前 R 个主要的模式具有较小或零的虚部和较大的实部, 体现了信号宏观上的趋势变

化。最后一列我们可以看到特征值的虚部在 $(-0.8, 0.8)$ 中均匀分布，这些特征值对应的模式具有分布较广的周期，没有特征周期。我们的数据都是实数，因此虚部在虚轴两侧是对称的，特征值都为共轭对。

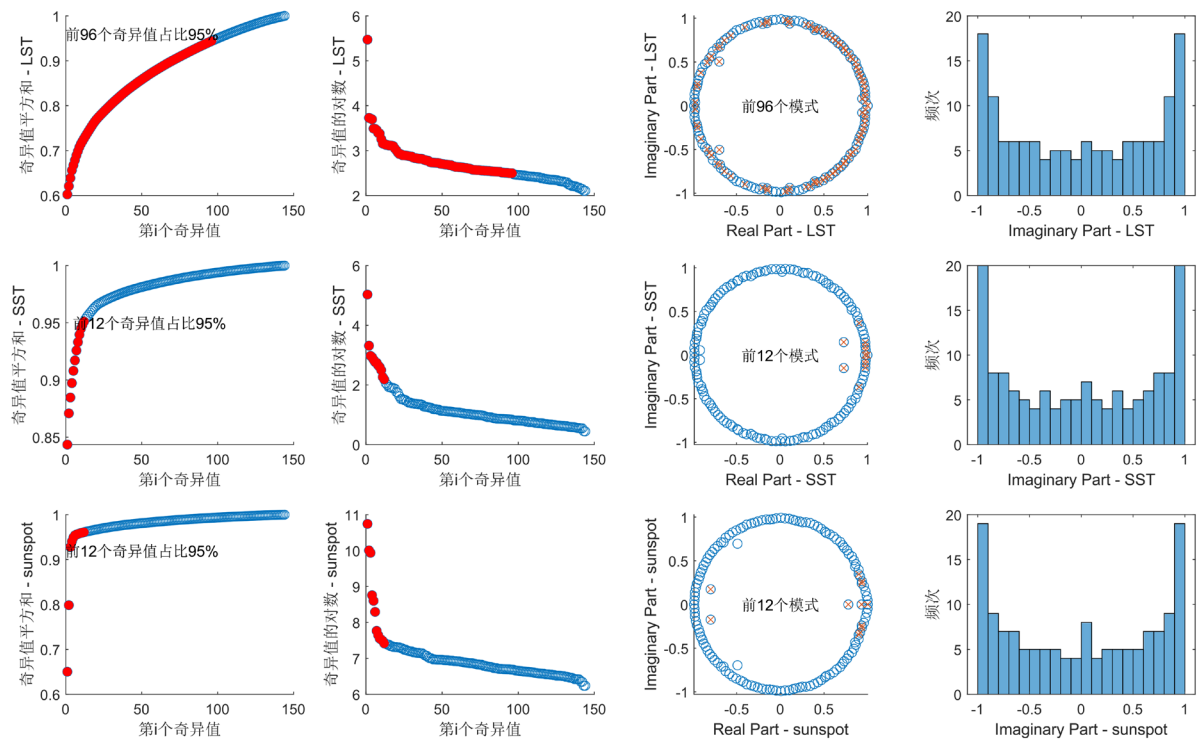


Figure 5. Distributions of singular values and eigenvalues of DMD linear operator from three monthly signals

图 5. 三种月度信号的奇异值和 DMD 线性算子的特征值分布

Table 4. The eigenvalues and periods of major modes from three monthly signals

表 4. 三种月度信号的主要模式的特征值和周期

特征值：三种信号的主要模式对应的特征值和周期	
<i>LST</i>	$0.94 \pm 0.08i$, $0.40 \pm 0.86i$, $0.70 \pm 0.50i$, $0.97 \pm 0.10i$, 1.00 , $0.40 \pm 0.89i$, $0.98 \pm 0.04i$;
主要周期	1 年, 1 个月, 2 个月, 长期趋势, 2 年
<i>SST</i>	$0.73 \pm 0.15i$, 0.97 , 1.00 , $0.98 \pm 0.06i$, $0.98 \pm 0.10i$, $0.91 \pm 0.37i$, $0.98 \pm 0.14i$
主要周期	6 个月, 长期趋势, 1 年, 3 个月
<i>sunspot</i>	0.78 , $0.93 \pm 0.24i$, 0.93 , 1.00 , $0.90 \pm 0.35i$, $0.93 \pm 0.27i$, $0.80 \pm 0.17i$, 0.91
主要周期	长期趋势, 4 个月, 3 个月, 6 个月

三种月度信号的最大特征值虚部很小，且 *SST* 和太阳黑子的前少数几个最大的特征值具有较小的虚部，意味着前几个主导模式为慢模式，有着较大的周期。*LST* 主要模式较多，表 4 展示了前 13 个特征值。贡献最大的一对共轭特征值体现了其显著的一年周期的特点： $1/0.08 \approx 12$ ；第二对和第六对特征值的实部较小而虚部较大，反映了快速振荡占据 *LST* 很大一部分特征；第四和第七对特征值的虚部对应 1 年和 2 年左右的周期。*SST* 的第一和第七对特征值反映了半年周期；第四和第五对特征值对应一年左右的周期；第六对特征值对应三个月左右周期。太阳黑子的前几对特征值反映了 3~6 个月的周期，贡献最大的特征值不包含虚部，体现了太阳黑子具有很强的趋势性。除了 *LST*，前少数几个特征值都对应较长的周期和长期趋势，意味着当 r 很低时，提取出来的模式很慢；当 r 增大时，特征值的虚部会越来越大，意味着

我们可以提取出更快的模式。图 6 展示了四个原始信号的几种典型的快模式和慢模式。当 r 接近 1 时，我们得到一个非常慢的模式，随着 r 增大，我们能够提取出包含多种尺度的模式。

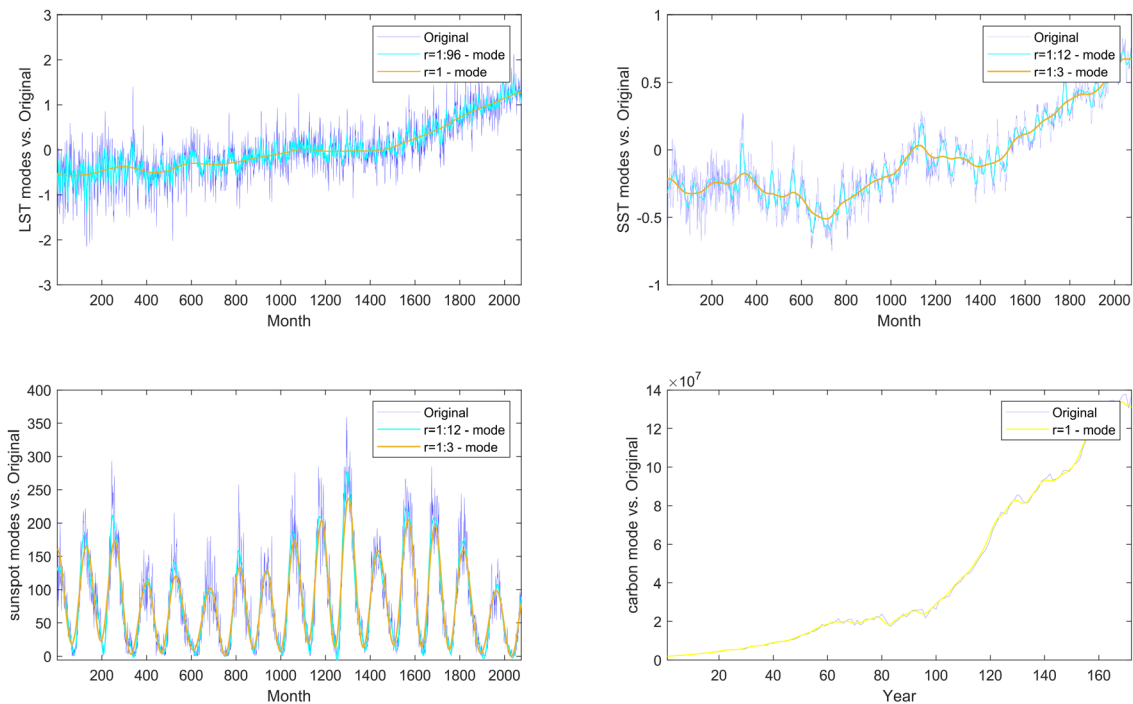


Figure 6. Several typical fast and slow modes
图 6. 几种典型的快慢模式

我们计算了各种模式之间的转移熵，它们代表着 8 种因果关系在不同模式之间的分布。为了确保计算出来的 TE 值是可信的，每种模式都被打乱并计算 20 次 TE 值，如果存在多于一个的打乱 TE 值大于原始顺序的 TE 值，则舍弃这个 TE 值，这等价于 95% 的可信度。我们在尝试中发现即使调整这个阈值或改变打乱次数，也不会对 TE 值在各种模式下的分布形状产生任何影响。即：不同的统计显著性不会改变 TE 值的分布，这体现了因果关系在不同模式之间的固有特征。图 7 展示了这些关联如何分布在各种模式中，表 5 为因果强度较大时对应的模式。总体上，关联集中于少数特征提取的慢模式，大部分模式之间不显著存在关联。由第一个特征值提取的主导模式作为因，它会对来自同一个气候对象中的其他较慢和稍快的模式产生显著影响，对其他气候对象的主导模式也有很大贡献，反过来则不成立。 LST 的最慢的模式约为一年周期，它对自身的一年和一两个月周期产生了显著影响，对 SST 的半年周期以及碳排放的长期趋势亦有很大贡献。注意到由前 11~13 个特征值提取的 LST 模式中包含了 1~12 个月和两年的周期，这些模式对自身的各种模式有强烈影响，但与内部其他模式和其他气候信号的模式之间关联极少，体现了 LST 中的快速振荡或噪声基本只影响自身的特点。 SST 模式中半年周期的驱动力极强，它对自身的三个月，半年，一年周期和长期趋势都有着显著影响；对 LST 的一年周期和碳排放亦有很大贡献。太阳黑子的最大主导模式，即长期趋势对 SST 的 6 个月以上周期有极大的贡献；前 9 个特征值下提取的模式仅包含长期趋势和 3~4 个月周期，这些模式的内部关联非常显著；其四个月周期和长期趋势对 LST 在一两个月和一年的尺度上有贡献，但偏弱；对碳排放的影响没有计算出任何有效的 TE 值。碳排放对 SST 的半年周期有强烈影响，对 LST 的一个月及一年周期有显著贡献，在其他尺度上的影响偏弱，气温对碳排放的影响也偏弱。

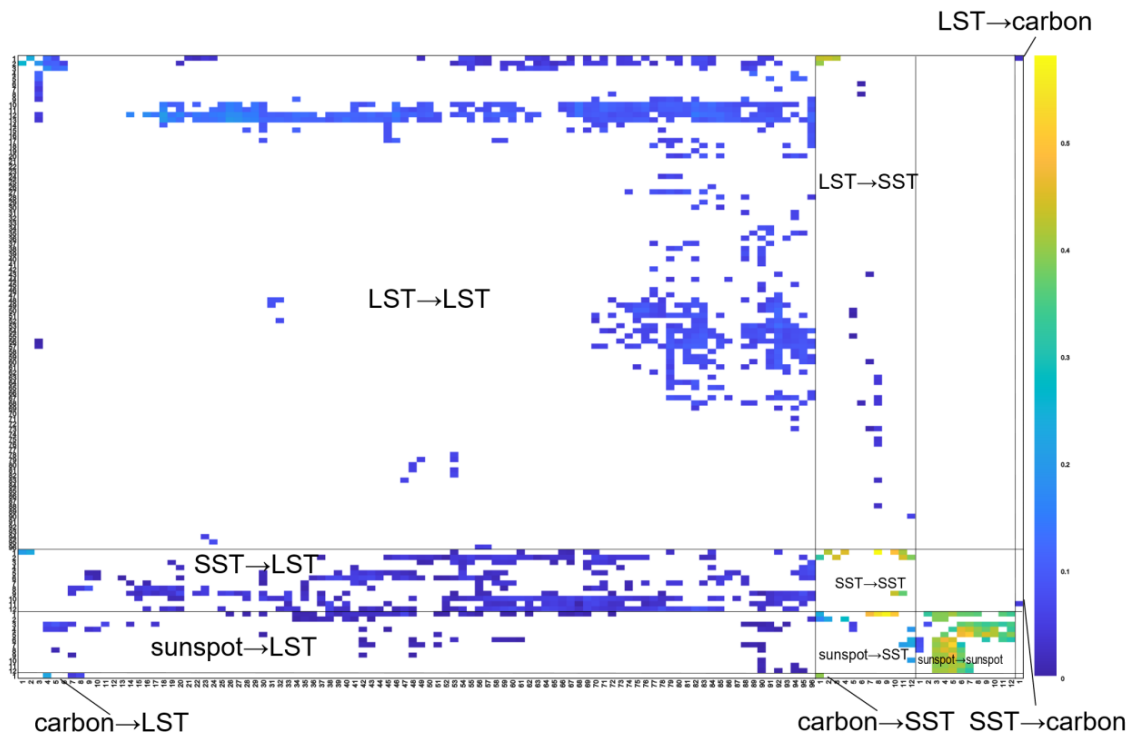


Figure 7. Causal intensity among different modes. The horizontal axis represents the cause, while the vertical axis represents the effect. From left to right, and top to bottom, they respectively represent *LST*, *SST*, *sunspots*, and *carbon emissions*. The color bar on the right indicates the magnitude of the TE values, where white regions in the plot indicate insignificant or zero TE values. The numbers on the horizontal and vertical axes represent the modes extracted by a certain number of eigenvalues for a specific climate variable

图 7. 各种模式之间的关联强度。横轴代表因，纵轴代表果，从左到右，从上到下分别为 *LST*，*SST*，太阳黑子和碳排放。右边颜色条为 TE 值的大小，图中白色区域表示 TE 值不显著或为零。横纵轴上的数字表示对某个气候对象由前若干个特征值提取的模式

Table 5. Modes corresponding to causal intensity around peak values

表 5. 因果强度在峰值附近对应的模式

模式对: $TE(U^m \rightarrow V^n)$ 较大时对应的模式, 记为 $m \rightarrow n$: TE	
<i>LST</i> → <i>LST</i>	2→1: 0.26, 1→2: 0.24, 12→18: 0.21, 13→18: 0.19, 2→3: 0.19, 3→4: 0.17
<i>SST</i> → <i>LST</i>	1→2: 0.23, 1→1: 0.21
<i>Sunspot</i> → <i>LST</i>	3→5: 0.09, 3→6: 0.07, 3→4: 0.06
<i>Carbon</i> → <i>LST</i>	1→4: 0.23, 1→8: 0.07, 1→3: 0.05
<i>LST</i> → <i>SST</i>	1→2: 0.44, 1→1: 0.43, 1→3: 0.42, 2→1: 0.40
<i>SST</i> → <i>SST</i>	1→8: 0.58, 1→10: 0.49, 1→4: 0.45, 2→3: 0.45, 2→11: 0.44, 9→10: 0.43, 1→2: 0.42, 1→11: 0.41
<i>sunspot</i> → <i>SST</i>	1→9: 0.57, 1→8: 0.54, 1→10: 0.50
<i>carbon</i> → <i>SST</i>	1→1: 0.40
<i>sunspot</i> → <i>sunspot</i>	7→5: 0.52, 5→7: 0.51, 9→5: 0.48, 6→4: 0.47, 5→6: 0.47, 4→6: 0.46
<i>LST</i> → <i>carbon</i>	1→1: 0.03
<i>SST</i> → <i>carbon</i>	11→1: 0.05

我们以各个模式之间最大的 TE 值作为因果强度的特征, 并把它们作为信息的传递写在图 8 的因果信息框架中。若不考虑括号内的自关联, *LST* 的净 TE 流入为 0.08, *SST* 的净 TE 流入为 1.13, 太阳黑子的净 TE 流出为 0.66, 碳排放的净 TE 流出为 0.55。总体上净 TE 流入和流出相等, 表示由四种气候对象建立的因果信息框架达到了一种平衡。

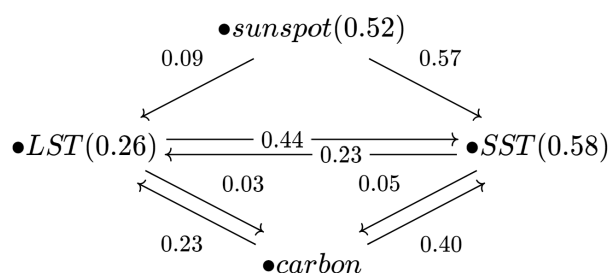


Figure 8. The causal information framework established by four climate variables, where the numbers on the arrows represent the maximum TE values between modes from different climate variables, and the numbers in parentheses represent the maximum TE values of modes within a particular climate variable

图 8. 四种气候对象建立的因果信息框架, 箭头上的数字为不同气候对象的模式之间的最大 TE 值, 括号内的数字为某个气候对象内部的模式间的最大 TE 值

4. 结论

我们的方法得到了因果关系在各种模式下的分布。关联集中于少数特征提取的慢模式。最慢的模式作为因, 它会对较慢和稍快的模式产生显著影响。太阳黑子内部模式之间的关联集中于长期趋势和 3~4 个月的周期, 这些尺度对 SST 的各种周期和长期趋势的贡献极强, 对 LST 在各种层面上的影响偏弱, 我们推断 SST 对太阳黑子的响应比 LST 更大。碳排放对 SST 的半年周期具有很大影响, 对 LST 的一年周期也有较大影响, 但在更快尺度上的影响偏弱, 气温反过来对碳排放的影响偏弱。LST 的一年周期对 SST 的半年周期和碳排放有较大贡献, LST 内部模式间的关联从快速振荡到长期趋势, 包含多种尺度。SST 的半年周期对自身, LST 的一年周期和碳排放均有很大贡献。两种气温之间的关联也聚集在半年和一年周期, LST 对 SST 的影响更大。LST, SST, 太阳黑子和碳排放四个对象构成的因果框架展示了它们之间信息的传递, 总体上看信息在这个框架中的流动是平衡的。在整个 1850~2022 的时间尺度上, SST 比 LST 吸收了更多的信息, 太阳黑子对气温的贡献略大于碳排放。

我们提出的 DMD 框架有着广阔的应用前景。第一, 多分辨率的能力。本文将 DMD 应用于整体时间尺度, 若切分成若干个时间段并用 DMD 分段处理, 能够分析某一时间段上的关联, 从而更深入地理解系统的演化特征。第二, 多因子分析的能力。影响地球气候的其他因素如火山活动, 气溶胶, 其他温室气体乃至太阳系中其他的天体活动都会对气候产生不可忽视的影响[12] [23], 即使这些序列的记录较少, DMD 仍然有着从中提取合适模式并分析因果关系的潜力。第三, 多数据类型的适用性。气候或其他复杂系统不仅由时间序列反映演化行为, DMD 可以应用于遥感影像等图像类型的数据[24]。第四, 预测的潜力。通过配合控制或一些机器学习方法, DMD 在最近的研究中也展现出其预测时间序列的潜力[25]。

参考文献

- [1] Schmid, P.J. (2010) Dynamic Mode Decomposition of Numerical and Experimental Data. *Journal of Fluid Mechanics*, **656**, 5-28. <https://doi.org/10.1017/S0022112010001217>
- [2] Grosek, J. and Nathan Kutz, J. (2014) Dynamic Mode Decomposition for Real-Time Background/Foreground Separation in Video.
- [3] Tirunagari, S., et al. (2015) Detection of Face Spoofing Using Visual Dynamics. *IEEE Transactions on Information Forensics and Security*, **10**, 762-777. <https://doi.org/10.1109/TIFS.2015.2406533>
- [4] Berger, E., et al. (2014) Dynamic Mode Decomposition for Perturbation Estimation in Human Robot Interaction. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, Edinburgh, 25-29 August 2014, 593-600. <https://doi.org/10.1109/ROMAN.2014.6926317>

- [5] Brunton, B.W., *et al.* (2016) Extracting Spatial-Temporal Coherent Patterns in Large-Scale Neural Recordings Using Dynamic Mode Decomposition. *Journal of Neuroscience Methods*, **258**, 1-15. <https://doi.org/10.1016/j.jneumeth.2015.10.010>
- [6] Tirunagari, S., Kouchaki, S., Poh, N., Bober, M. and Windridge, D. (2017) Dynamic Mode Decomposition for Univariate Time Series: Analysing Trends and Forecasting.
- [7] Gray, L.J., Beer, J., Geller, M., Haigh, J.D., Lockwood, M., Matthes, K. and White, W. (2010) Solar Influences on Climate. *Reviews of Geophysics*, **48**, RG4001. <https://doi.org/10.1029/2009RG000282>
- [8] Pangburn, D. (2014) Influence of Sunspots on Global Mean Surface Temperature. *Energy & Environment*, **25**, 1455-1471. <https://doi.org/10.1260/0958-305X.25.8.1455>
- [9] Scafetta, N. (2014) Global Temperatures and Sunspot Numbers. Are They Related? Yes, but Non Linearly. A Reply to Gil-Alana *et al.* (2014). *Physica A: Statistical Mechanics and Its Applications*, **413**, 329-342. <https://doi.org/10.1016/j.physa.2014.06.047>
- [10] Kristoufek, L. (2017) Has Global Warming Modified the Relationship between Sunspot Numbers and Global Temperatures? *Physica A: Statistical Mechanics and Its Applications*, **468**, 351-358. <https://doi.org/10.1016/j.physa.2016.10.089>
- [11] Knuth, K.H., Gotera, A., Curry, C.T., Huyser, K.A., Wheeler, K.R. and Rossow, W.B. (2013) Revealing Relationships among Relevant Climate Variables with Information Theory.
- [12] Bhaskar, A., Ramesh, D.S., Vichare, G., Koganti, T. and Gurubaran, S. (2017) Quantitative Assessment of Drivers of Recent Global Temperature Variability: An Information Theoretic Approach. *Climate Dynamics*, **49**, 3877-3886. <https://doi.org/10.1007/s00382-017-3549-5>
- [13] Papan, A., Kugiumtzis, D. and Larsson, P.G. (2012) Detection of Direct Causal Effects and Application to Epileptic Electroencephalogram Analysis. *International Journal of Bifurcation and Chaos*, **22**, Article ID: 1250222. <https://doi.org/10.1142/S0218127412502227>
- [14] Zhao, P. and Lai, L. (2022). Analysis of Knn Density Estimation. *IEEE Transactions on Information Theory*, **68**, 7971-7995. <https://doi.org/10.1109/TIT.2022.3195870>
- [15] Estimated Global Land-Surface TAVG Based on the Complete Berkeley Dataset. <https://berkeleyearth.org/data/>
- [16] Kennedy, J.J., Rayner, N.A., Atkinson, C.P. and Killick, R.E. (2019) An Ensemble Data Set of Sea-Surface Temperature Change from 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set. *Journal of Geophysical Research: Atmospheres*, **124**, 7719-7763. <https://doi.org/10.1029/2018JD029867>
- [17] SILSO, World Data Center-Sunspot Number and Long-Term Solar Observations, Royal Observatory of Belgium, On-Line Sunspot Number Catalogue.
- [18] Global Carbon Project (2022) Supplemental Data of Global Carbon Budget 2022 (Version 1.0) [Data Set]. Global Carbon Project.
- [19] Tu, J.H. (2013) Dynamic Mode Decomposition: Theory and Applications. Doctoral Dissertation, Princeton University, Princeton.
- [20] Kutz, J.N., Brunton, S.L., Brunton, B.W. and Proctor, J.L. (2016) Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems. Society for Industrial and Applied Mathematics, Philadelphia. <https://doi.org/10.1137/1.9781611974508>
- [21] Kraskov, A., Stögbauer, H. and Grassberger, P. (2004) Estimating Mutual Information. *Physical Review E*, **69**, Article ID: 066138. <https://doi.org/10.1103/PhysRevE.69.066138>
- [22] Shi, J.F., Chen, L.N. and Aihara, K. (2022) Supplementary Information for “Embedding Entropy: A Nonlinear Measure of Dynamical Causality”.
- [23] Scafetta, N. (2010) Empirical Evidence for a Celestial Origin of the Climate Oscillations and Its Implications. *Journal of Atmospheric and Solar-Terrestrial Physics*, **72**, 951-970. <https://doi.org/10.1016/j.jastp.2010.04.015>
- [24] Kutz, J.N., Fu, X. and Brunton, S.L. (2016) Multiresolution Dynamic Mode Decomposition. *SIAM Journal on Applied Dynamical Systems*, **15**, 713-735. <https://doi.org/10.1137/15M1023543>
- [25] Mansouri, A., Abolmasoumi, A.H. and Ghadimi, A.A. (2023) Weather Sensitive Short Term Load Forecasting Using Dynamic Mode Decomposition with Control. *Electric Power Systems Research*, **221**, Article ID: 109387. <https://doi.org/10.1016/j.eprsr.2023.109387>