Sapienza
Università di Roma

Detecting Causal Interactions between Atlantic and Pacific Ocean Variability

Facoltà di Ingegneria dell'Informazione, Informatica e Statistica
Laurea Magistrale in Data Science

Jonas Barth
ID number 2050678

Supervisor	Co-Supervisor
Prof. Paolo Tieri	Dott. Alessio Bellucci Dott. David Docquier Dott. Valerio Lembo Prof. Manuela Petti

Academic Year 2024/2025

Thesis defended on 30 May 2025
in front of a Board of Examiners composed by:

Prof. Casalicchio Emiliano (chairman)
Prof. Colonnese Stefania
Prof. Maggi Bernardo
Prof. Piccialli Veronica
Prof. Polito Pompeo
Prof. Quattrociocchi Walter
Prof. Tieri Paolo

Detecting Causal Interactions between Atlantic and Pacific Ocean Variability
Tesi di Laurea Magistrale. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: jonas.barth.95@gmail.com

Dedicated to my family and friends. I especially want to thank my parents, who have supported me in all of my choices with their time, love, and money, always believing in my ability to achieve my goals. Next, I want to thank my brothers, both older and younger, who have been and continue to be role models for me in life and shown me that starting a new life abroad is an exciting opportunity not to be missed. I also dedicate this thesis to my friends here in Rome, who have made this journey not only more bearable but also genuinely enjoyable. Your friendship has often been the perfect distraction from the pressures of university life, reminding me to laugh, explore, and appreciate the present. Finally, I want to thank my girlfriend Angelica for giving me the idea to study in Italy and for helping me navigate the difficulties of living here.

Abstract

This thesis investigates the causal interactions between Atlantic and Pacific Sea Surface Temperature (SST) variability. Specifically, we analyse the Atlantic Multidecadal Variability (AMV) and Pacific Decadal Oscillation (PDO), two significant modes of climate variability that influence global weather patterns and climate dynamics, as well as first principal components in the Pacific. Using the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST) from the Met Office and the Extended Reconstructed SST Version 5 (ERSSTV5) dataset from the National Oceanic and Atmospheric Administration (NOAA), the study calculates the AMV and PDO indices as well as first principal components in the Pacific, and applies three causal discovery methods - Granger Causality, Peter and Clark Momentary Conditional Independence Plus (PCMCI+), and Liang Kleeman Information Flow (LKIF) - to uncover causal relationships in different settings.

The results reveal that all three methods identify causal links from the Atlantic to the Pacific Ocean, with Granger Causality and LKIF consistently detecting connections from the AMV to the Pacific. PCMCI+, however, only finds such links when considering three principal components alongside the PDO. The study highlights that the connections from AMV to PDO are is weaker than connections from AMV to other Pacific components and that these Pacific components have a stronger link to the PDO than the AMV. This suggests that other components such as the tropical Pacific, may act as mediators in the influence of the AMV on the PDO.

This research contributes to the existing literature by providing a comparative analysis of different causal discovery methods and offering insights into the complex interactions between the AMV and PDO. The findings underscore the importance of considering multiple components and indirect pathways in understanding the causal dynamics of climate variability and to employ several causal methods to account for uncertainties that come with each method. Future work could explore the subdivision of the AMV into tropical and subtropical components and incorporate additional climate indices to further refine the causal models.

Contents

Chapter 1

Introduction

The year 2024 was the hottest on record and the first to reach an average global temperature 1.5°C above pre-industrial levels. Notably, the ten warmest years in recorded history have all occurred within the past decade [170, 7, 21, 104, 11]. The rising global mean surface temperature as seen in Figure 1.1 has many implications. The previous year saw concentrations of greenhouse gases such as carbon dioxide (CO2) and methane (CH4) reaching record highs, and upper Ocean heat content also rising to unprecedented levels. As a result of the warming, land and sea surface temperatures were a record of 2.3°C and 1.1°C warmer compared to pre-industrial levels respectively, cumulative ice loss from glaciers and ice sheets was the largest in recorded history, and sea levels have risen to a new record high [171, 58]. An alarming consequence of this warming trend is the increasing frequency and intensity of extreme weather events, including heavy rainfall leading to devastating floods, prolonged heatwaves, severe droughts, and powerful storms. Figure 1.2 shows some of these extreme climate events from 2024.

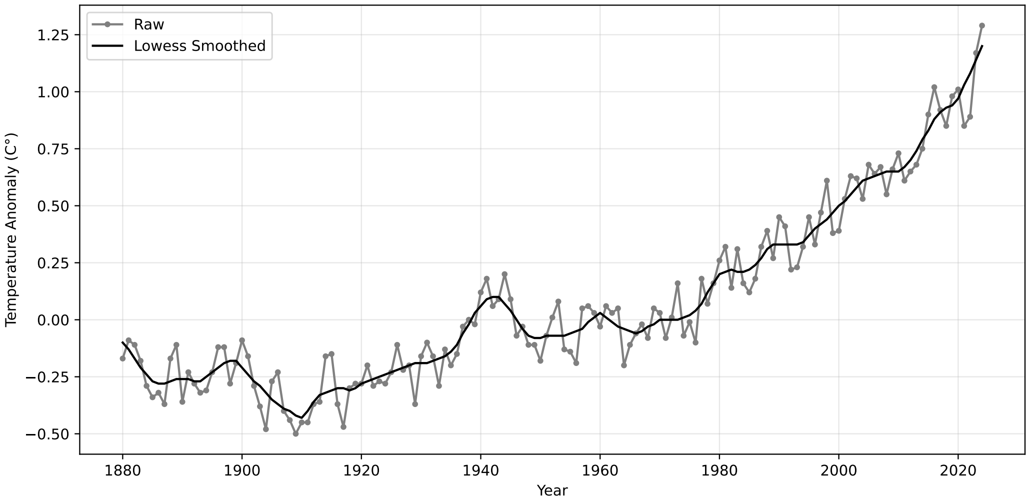
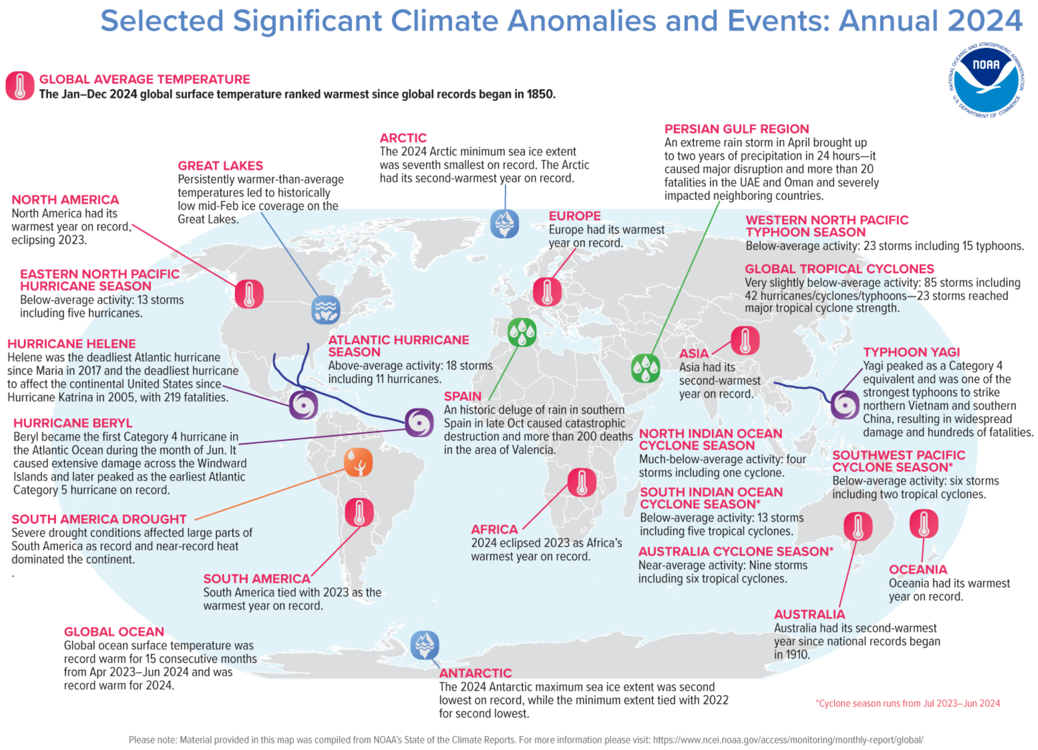


Figure 1.1 Change in global mean surface temperature from 1880 to 2024 relative to the average over the 1951-1980 period [101].

Studying the climate is paramount to understanding how climate change will impact the Earth in the future and how we as humans will have to mitigate and to adapt to its impacts. The Earth's climate system is complex and interconnected, influencing everything from weather patterns and ocean currents to agricultural productivity and biodiversity. In addition to disrupting natural ecosystems, climate change also pose significant risks to human health, infrastructure, and global economies. It exacerbates resource scarcity, forces communities to migrate due to uninhabitable conditions, and increases the risk of conflict over dwindling resources such as fresh water and arable land [62].



Selected Significant Climate Anomalies and Events: Annual 2024

GLOBAL AVERAGE TEMPERATURE
The Jan–Dec 2024 global surface temperature ranked warmest since global records began in 1850.

NORTH AMERICA
North America had its warmest year on record, eclipsing 2023.

GREAT LAKES
Persistently warmer-than-average temperatures led to historically low mid-Feb ice coverage on the Great Lakes.

ARCTIC
The 2024 Arctic minimum sea ice extent was seventh smallest on record. The Arctic had its second-warmest year on record.

PERSIAN GULF REGION
An extreme rain storm in April brought up to two years of precipitation in 24 hours—it caused major disruption and more than 20 fatalities in the UAE and Oman and severely impacted neighboring countries.

EUROPE
Europe had its warmest year on record.

WESTERN NORTH PACIFIC TYPHOON SEASON
Below-average activity: 23 storms including 15 typhoons.

EASTERN NORTH PACIFIC HURRICANE SEASON
Below-average activity: 13 storms including five hurricanes.

GLOBAL TROPICAL CYCLONES
Very slightly below-average activity: 85 storms including 42 hurricanes/cyclones/typhoons—23 storms reached major tropical cyclone strength.

HURRICANE HELENE
Helene was the deadliest Atlantic hurricane since Maria in 2017 and the deadliest hurricane to affect the continental United States since Hurricane Katrina in 2005, with 219 fatalities.

ATLANTIC HURRICANE SEASON
Above-average activity: 18 storms including 11 hurricanes.

ASIA
Asia had its second-warmest year on record.

TYPHOON YAGI
Yagi peaked as a Category 4 equivalent and was one of the strongest typhoons to strike northern Vietnam and southern China, resulting in widespread damage and hundreds of fatalities.

SPAIN
An historic deluge of rain in southern Spain in late Oct caused catastrophic destruction and more than 200 deaths in the area of Valencia.

HURRICANE BERYL
Beryl became the first Category 4 hurricane in the Atlantic Ocean during the month of Jun. It caused extensive damage across the Windward Islands and later peaked as the earliest Atlantic Category 5 hurricane on record.

NORTH INDIAN OCEAN CYCLONE SEASON
Much-below-average activity: four storms including one cyclone.

SOUTHWEST PACIFIC CYCLONE SEASON*
Below-average activity: six storms including two tropical cyclones.

SOUTH INDIAN OCEAN CYCLONE SEASON*
Below-average activity: 13 storms including five tropical cyclones.

SOUTH AMERICA DROUGHT
Severe drought conditions affected large parts of South America as record and near-record heat dominated the continent.

AFRICA
2024 eclipsed 2023 as Africa's warmest year on record.

AUSTRALIA CYCLONE SEASON*
Near-average activity: Nine storms including six tropical cyclones.

OCEANIA
Oceania had its warmest year on record.

SOUTH AMERICA
South America tied with 2023 as the warmest year on record.

AUSTRALIA
Australia had its second-warmest year since national records began in 1910.

GLOBAL OCEAN
Global ocean surface temperature was record warm for 15 consecutive months from Apr 2023–Jun 2024 and was record warm for 2024.

ANTARCTIC
The 2024 Antarctic maximum sea ice extent was second lowest on record, while the minimum extent tied with 2022 for second lowest.

*Cyclone season runs from Jul 2023–Jun 2024

Please note: Material provided in this map was compiled from NOAA's State of the Climate Reports. For more information please visit: https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/

Figure 1.2 World map annotated with significant climate events in 2024 [104].

One of the most important variables in the Earth's climate system is Sea Surface Temperature (SST). Oceans cover approximately 70% of the planet's surface, playing a crucial role in shaping both small and large-scale weather patterns by transferring moisture and heat to the atmosphere [114]. Additionally, the oceans have absorbed about 90% of the excess heat generated by human emissions to date [47, 64]. Changes in ocean temperature significantly impact ocean circulation patterns, which are responsible for distributing both cold and warm water around the globe. As SST rises, it also leads to an increase in atmospheric water vapor, which in turn influences precipitation systems, raising the likelihood of heavier rainfall and snowfall. This shift can also alter storm patterns, heightening the risk of severe storms or prolonged droughts [156]. Hence, studying SST is essential for understanding, quantifying, and predicting future climate dynamics, as it helps us assess potential regional and global impacts.

The Atlantic Multidecadal Variability (AMV) [72] and the Pacific Decadal Oscillation (PDO) [93] are two significant modes of climate variability that are closely linked to fluctuations in SST. These oceanic patterns, driven by changes in SST in the Atlantic and Pacific Oceans respectively, have significant implications for global weather systems, including shifts in rainfall patterns, storm intensity, and temperature extremes. By investigating the links between SST in both the Atlantic and Pacific regions, we gain insight into the natural climate cycles that influence multidecadal weather variations, as well as the interaction between these cycles and anthropogenic climate change. Moreover, we gain a better understanding of how the Atlantic and Pacific Oceans influence each other, which is not well understood as of now. Previous research has shown a link between the AMV and PDO [179, 178, 135], that we seek to further investigate and complement with this work. The interactions between variables in the Earth's climate system represent complex dynamical systems, which are difficult to study.

Traditionally, causal relationships are investigated with interventional experiments - think of Galileo dropping a cannon balls from the tower of Pisa [37], Mendel cross breeding peas [38], and James Lind finding a remedy for scurvy by supplementing sailors' diet with lemons [155]. These are the gold standard in medicine and social sciences, however on the scale of the climate system, these are technically infeasible and ethically questionable to say the least, and so are out of the question. An alternative is to use computer simulations, where with the help of experts we construct a physical model of the system we want to study. While this does not require to play with the climate directly, this involves a detailed knowledge of the climate system as well as access to substantial computing resources. Moreover, simply because an effect is observed in a model, does not mean that we uncover the causal relationships between variables in the system, as model results may arise from several influences. Hence it is necessary to rely on other methods to reconstruct causal networks from timeseries observations, methods that are driven by the data itself. This is where causal methods come into the picture. Thanks to the rapid growth of available data in the Earth Science field, it is now possible to apply these data-driven methods to infer causal dependencies between timeseries. In this thesis we choose Granger Causality [53], Peter and Clark Momentary Conditional Independence + (PCMCI+) [131], and Liang Kleeman Information Flow (LKIF) [88] (explained in Chapter 3.5.2) as causal methods. The aim of this thesis is to apply three causal discovery methods to progressively more complex systems: (1) the AMV and PDO indices; (2) the AMV, PDO, and the first three varimax-rotated principal components derived from extrapolar Pacific SSTs; and (3) the AMV, PDO, and the first five such components. By analyzing these systems, the goal is to identify recurring patterns across the linkages found between the Atlantic and Pacific Oceans. Using multiple methods provides a broader basis for comparison, as each approach has its own strengths and limitations and may yield different causal structures. Thus, it also takes into account the uncertainties that come with each method and consistent patterns emerging across methods may indicate robust, physically meaningful relationships within the climate system.

1.1 Climate Indices

Our climate is an extremely complex system, with countless interacting processes influencing its overall state. Monitoring such a complicate system by looking at the raw data is nearly impossible, hence we would like a simplified view that summarises parts of the climate system that we are interested in. For example, describing the atmospheric flow at all longitudes, latitudes, and altitudes on Earth would be a difficult task. A climate index is a a summary of an aspect of the climate system, providing a simplified overview of a feature of the system in question [103, 168]. The aspect may consist of one or several geophysical variables, depending on the desired quantity to measure. Indices help to predict weather, track climate variability, and detect climate change.

One of the most widely known indices is the El Niño Southern Oscillation (ENSO), which regularly makes an appearance in mainstream media [48, 3, 56, 46, 145]. ENSO is a phenomenon in the equatorial eastern Pacific, on the west coast of South America, with local effects on marine life and global impacts the climate and on weather patterns, affecting agriculture, storm frequency and magnitude, and precipitation among other things. Under normal conditions, so called westward blowing trade-winds move warm water from South America towards Asia, causing colder water from deeper depths to rise up. These normal conditions are upset by the El Niño and La El Niña phases respectively. During the warm El Niño phase, the trade winds weaken and push warm water to the Americas' west coast. This in turn reduces atlantic hurricane and tropical cyclone activity, causes drier than usual conditions in Southeast Asia and the Pacific Northwest, and wetter than usual conditions in coastal South America and Southern US states. During the cold La El Niña phase, stronger trade-winds push more warm surface water towards Asia, bringing colder water to the surface. As a result, Southeast Asia and the Pacific Northwest experience wetter than usual conditions whereas coastal South America and the Southern US suffer from droughts. Additionally, it can also lead to a more severe hurricane season, stronger monsoons and typhoons [107, 160, 105, 106]. While ENSO has a global impact on the climate, it is mainly a tropical phenomenon. In this study, we are more interested in the connections between the North Atlantic and North Pacific basins, which is not captured by ENSO.

1.1.1 AMV

The Atlantic Multidecadal Variability (AMV) is a low-frequency variation of the spatially averaged SST anomalies (SSTa) in the North Atlantic basin. The AMV index switches between warm and cold phases every 20-40 years, completing one warm-cold cycle in 60-80 years (Fig. 1.3). Originally defined by Kerr as the Atlantic Multidecadal Oscillation (AMO) in the year 2000 [72], some researches have sinced swapped the word *Oscillation* for *Variability*, since oscillation implies a regular cyclical pattern which is larger claim than using variability, which describes a broader range of SSTa fluctuations.

The Atlantic Multidecadal Variability (AMV) influences both regional and global climate, shaping precipitation patterns, land and sea surface temperatures, as well as storm tracks and intensities across multiple continents and ocean basins. In West

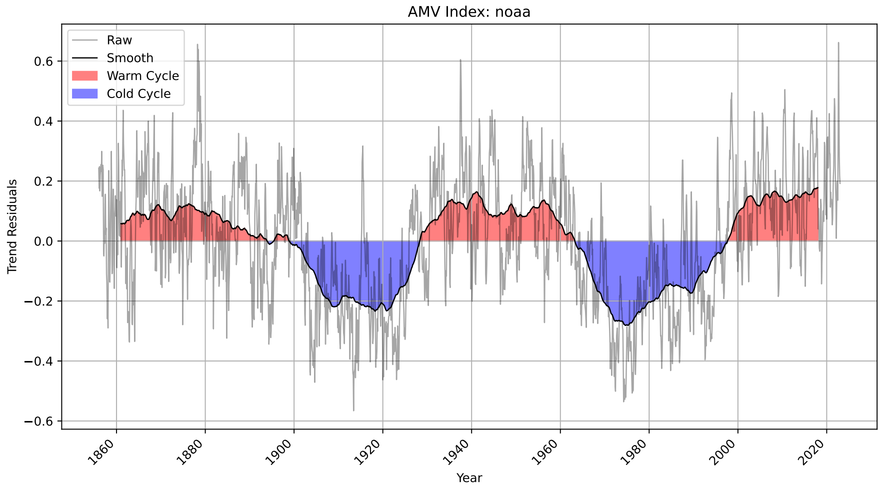


Figure 1.3 The raw and smoothed AMV index from the National Oceanic and Atmospheric Administration (NOAA) [112]. The raw index is calculated from the Kaplan dataset, over latitudes 0-70N, and smoothed with a 121 month running mean (black line). NOAA does not provide the longitudes or detrending technique used. Areas shaded in red represent positive phases of the index, blue areas the negative phases.

Africa, it has been linked to multi-decadal droughts through its modulation of the West African Monsoon variability [177]. During its warm phases, the AMV affects Eurasian winter temperatures, for instance, by weakening mid-latitude westerlies and contributing to winter cooling over parts of Eurasia [57, 17]. The AMV also plays a key role in tropical cyclone activity, enhancing hurricane frequency and intensity in the Atlantic and influencing Pacific cyclone variability [176, 74, 75, 180, 15]. In South America, it can amplify ENSO-driven rainfall anomalies, further affecting regional precipitation patterns [70]. The AMV dynamics are linked to both internal and external mechanisms. Internally, the Atlantic Meridional Overturning Circulation (AMOC) is thought to be a main driver [25]. It transports warm, surface waters from the tropics to higher latitudes and returns cooler, deeper waters toward the equator. This mechanism is particularly significant for the climate of regions such as Europe and North America, as it contributes to the regulation of temperature and weather patterns. The circulation is primarily driven by differences in seawater temperature and salinity - factors that influence water density. Warm water, being less dense, remains near the surface, while colder and saltier water, which is denser, sinks to greater depths. As warm water travels northward via the Gulf Stream, it gradually cools and increases in density, eventually sinking in the North Atlantic and flowing back to mid-latitudes. Additionally, during the formation of sea ice near the polar regions, freshwater is removed from the ocean, leaving behind higher salinity water. This process further increases the density of the remaining water, enhancing its downward movement [127, 110].

1.1.2 PDO

The PDO is an irregular sub-decadal to decadal pattern of monthly SST anomalies in the North Pacific basin. Like the AMV it has warm and cool phases, however the cycle is much less regular than that of the AMV. PDO cycles can last from years (cold phase from 1999-2002) to decades (warm phase 1977-1999) [102]. During the warm phase, the north eastern Pacific experiences warmer-than-normal SST, shown as warm horseshoe pattern in which we find a cold tongue from East Asia (see the left figure in fig 1.4). In contrast, during the cold or negative phase, the pattern is inversed. The PDO index is the first principal component obtained by applying Empirical Orthogonal Function (EOF) analysis (section 3.4.3), which goes by the name Principal Component Analysis (PCA) in the field of Computer Science, to SST anomalies in the North Pacific. The term PDO was first introduced by Mantua et al. in 1997 [93] as an irregular interannual to interdecadal pattern of ocean-atmosphere variability in the North Pacific basin. The observed phenomenon might also be referred to as an ENSO-like variability [181], however since Mantua's work, PDO is the term that is used.

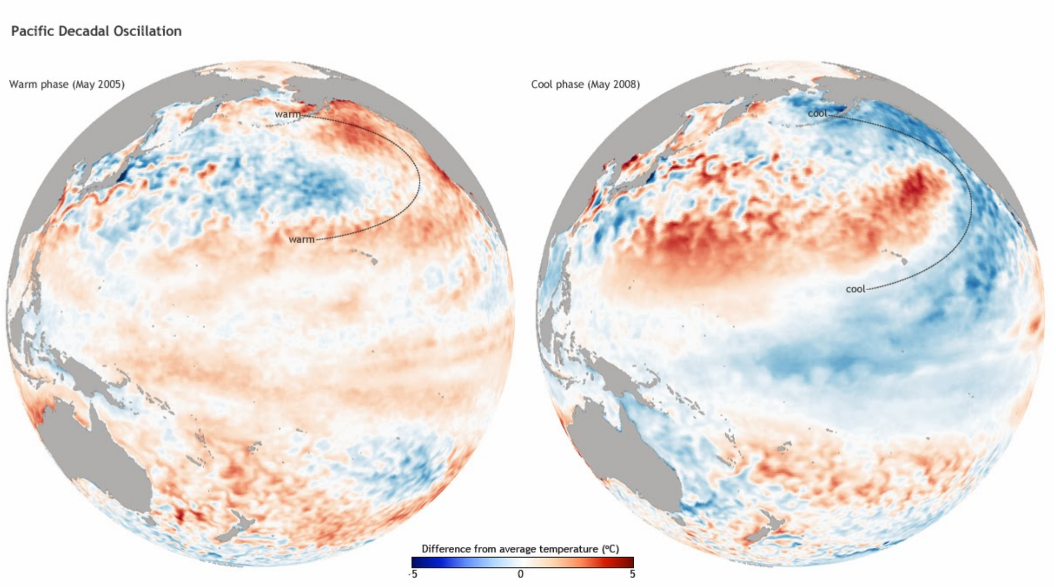


Figure 1.4 Warm and cold phases of the PDO. The North Pacific shows a distinct horseshoe shape at the west coast of the United States, wrapped around a tongue of colder/warmer temperatures from the West Pacific. Image [44].

The causes of the PDO are still actively being researched. Some say the PDO is just long lasting ENSO fingerprints [98]. Other possible causes are random variations in a pressure system near the Aleutian Islands in the Gulf of Alaska, referred to as the Aleution Low, atmospheric teleconnections from the tropics which alter near-surface air temperature, humidity, wind, and clouds in the North Pacific, or midlatitude ocean dynamics [108]. Many impacts of PDO are linked to ENSO. During the warm phases of the PDO, certain regions experience anomalously dry periods, including eastern Australia, Korea, Japan, the Russian Far East, interior Alaska, and a belt stretching from the Pacific Northwest to the Great Lakes and the

Ohio Valley. Additionally, much of Central America and northern South America also tend to be drier during these phases. Conversely, the warm PDO phases coincide with anomalously wet periods in the coastal Gulf of Alaska, the southwest US and Mexico, southeast Brazil, south central South America, and western Australia [91]. The PDO also modulates ENSO, amplifying the wet and dry phases when the two are in phase [162]. Furthermore, the PDO has been shown to influence marine heatwaves along the Northeast Pacific coast [129]. The negative phase of the PDO is associated with favorable conditions for tropical cyclone development in the mid-latitudes of East Asia [81]. Additionally, the PDO is linked to global extreme precipitation events [163] and plays a role in modulating the East Asian winter monsoons [73].

Chapter 2

Literature Review

The impact of the AMV on the PDO is an active area of research, with researchers applying a wide variety of methods to study their relationship and understanding the underlying causal graph as well as proposing different processes that govern the interaction between the two indices.

The most common way to study the relationship between AMV and PDO are model simulations. Overall, the connections found using model simulations are almost exclusively from the Atlantic to the Pacific. Using the global coupled ocean-atmosphere model CM2.1 from the Geophysical Fluid Dynamics Laboratory (GFDL), Delworth and Zhang [178] found that the AMO provides a source of multidecadal variability to the North Pacific, acting as a trigger to a regime change in North Pacific SST with a decade long lead time. Their proposed mechanism is that a positive AMO phase, changes North Pacific wind patterns, which shifts ocean currents in the Pacific, leading to warmer waters that push the Pacific into a state like the positive PDO. Okumura et al. [117] compare various simulation models and find oceanic and atmospheric pathways from the Atlantic to the North Pacific, where freshwater input to the North Atlantic transports colder water to the Arctic Ocean which ends up in the North Pacific via the Bering Strait. Others, like Kurcharski et al. [79], using the International Centre for Theoretical Physics AGCM (ICTPAGCM) model have found that the AMV has a substantial impact on equatorial pacific SST, thought to be modulated via the Walker Circulation, an atmospheric circulation bringing air from the Atlantic Tropics to the Pacific Tropics. The Walker Circulation as a modulator was also highlighted by other scholars. Using the NCAR-CESM1 model, Trascasa-Castro et al. [149] found that during the warm phase of the AMV, the Walker circulation impacts sea surface temperature patterns in the tropical Pacific, leading to weaker El Niño and La Niña phases. The impact of the AMV on ENSO via Walker Circulation was also studied by Levine et al. [83] who used the CM2.1 model to fix the AMV and let other regions adapt to it. Johnson et al. [66] set the Atlantic and Pacific basins to observational data and let the other global regions evolve using the CESM1.0 and MIROC3.2m climate models and found that Atlantic variability can influence the Pacific - and thus the PDO - through atmospheric teleconnections via both equatorial and north tropical pathways. Chen et al. [16] found that a negative AMV phase impacts the Arctic Oscillation (AO) that in turn changes the strength of ENSO via the Aleutian Low and and Polar Vortex.

Other studies instead calculate simpler statistical associations between observations to find relationships between the Atlantic and Pacific basin. d'Orgeville and Peltier [23] regress Pacific SSTs onto the AMO index at different time lags and find high spatial correlations with the PDO pattern, suggesting a link between the two on multidecadal timescales where the AMV leads the PDO by 13 years. Using correlations to define relationships between locations or events in climate data has also been used for discovering drought patterns in Europe [51], understanding South American rainfall dynamics [167, 10], predicting extreme floods in the Andes [9], and finding global rainfall teleconnections [8]. The issue with these simpler statistical approaches, is that they just tell us about statistical associations but do not identify causality, however it may be defined. Additionally, any relationship they find between variables is undirected and they struggle with common drivers and autocorrelated timeseries.

Causal methods on the other hand however, apply stricter constraints and allow us to go beyond simple statistical associations by providing directionality and intensity to causal links. While causal methods have already furthered the understanding of our climate, they have not seen widespread application to study the links between AMV and PDO. The goal of causal methods, is to detect robust causal links between observational variables, from which we can then reconstruct a causal graph. There is a large variety of causal methods, the approaches often differ on how to reconstruct the underlying network from the observations. Lately there has been an effort of introducing causality into Machine Learning (ML) [89, 5, 147, 45, 67]. Some examples include heatwave predictions using Graph Neural Networks (GNN) [84], making transformers causal [99, 109, 97], forecasting weather with GNNs [71, 80, 6] and transfomers [2], and finding links between solar wind and magnetosphere [90]. Like with regular ML, these methods tend to be data hungry, lack interpretability, and have no clear statistical significance baked in. Docquier et al. [34] applied two causal methods - Liang Kleeman Information Flow (LKIF) [88] and Peter and Clark Momentary Conditional Independence (PCMCI) [133] - to different climate indices, including AMV and PDO; they do not find any substantial influence between AMV and PDO, except from AMV to PDO with a lag of 8 months using the PCMCI method. Karmouche et al. [69] apply PCMCI+, a version of PCMCI that can find instantaneous links, to a system consisting of the AMV, PDO, Pacific North-American (PNA) pattern, and the Pacific South-American (PSA) pattern, and find a link from AMV to PDO with an 11 month lag. The influence from the AMV to the PDO is also identified by Vannitsem and Liang [157], who apply the Liang-Kleeman Information Flow (LKIF) method to a network of climate indices - including the AMV, PDO, ENSO, AO, and NAO - at time lags from 0 to 60 months. Their analysis reveals a directional and significant influence from the AMV to the PDO, but no significant impact from the PDO to the AMV. Silini et al. [140] used Pseudo Transfer Entropy (pTE), a linear approximation of the expensive to calculate Transfer Entropy (TE), and found links from AMV to ENSO as well as from PDO to AMV. In contrast to the findings by others, they do not find any links from AMV to PDO. The goal of this thesis is to complement the insight of the existing literature on the relationship between the AMV and PDO and to compare the results of **several** causal methods.

Chapter 3

Data

We calculate the AMV and PDO indices freely available world wide spatio-temporal SST data. SST is one of the first climate variables studied and remains one of the most important ones for tracking the Earth's climate system. Being the boundary between the ocean and atmosphere, SST helps us to understand the energy flow between the two and how they shape the weather and climate [22].

3.1 Data Structure

The data is spatio-temporal, where each data point corresponds to a location at a given point in time, on Earth. The entire dataset is a multidimensional array, indexed by two temporal dimensions, longitude ϕ and a latitude θ, and one temporal dimension t (Figure 3.1).

$$\mathbf{SST} \in \mathbb{R}^{t \times \phi \times \theta}$$

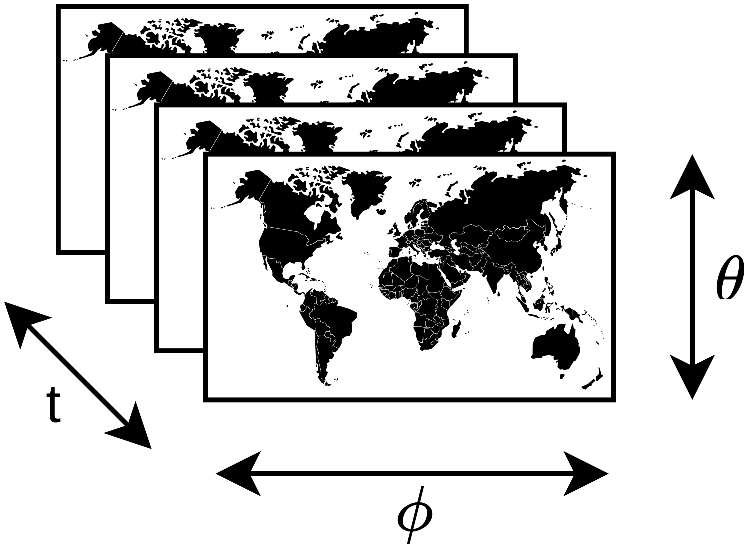


Figure 3.1 The SST through time t, with latitudes θ and longitudes ϕ. Each slice in the figure represent a point in time of the SST array $\mathbf{SST}(t_i, \phi, \theta)$.

3.2 Datasets

There is a variety of datasets that are freely available to download, with various temporal and spatial resolutions, as well as a range of time periods covered. The two dataset used in this thesis are the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) dataset from the Met Office in the UK, which goes back to 1871, and the NOAA Extended Reconstructed SST V5 (ERSSTV5) with data all the way to 1854. Table 3.1 shows a comparison of the main features of these two datasets.

Name	**Resolution** t	**Resolution** ϕ, θ	**Time Range**	**Source**
HadISST	Monthly	$\phi = [-180, 180], \theta = [-90, 90]$	1871 – present	[128]
ERSSTV5	Monthly	$\phi = [-180, 180], \theta = [-90, 90]$	1854 – present	[61]

Table 3.1 Summary of the HadISST and ERSSTV5 datasets used in this thesis.

It is important to note that the data points in these datasets are not the raw observations, but already a processed version, to account for measurement gaps, change in measurement techniques and tools. The earliest records were made from ships, where the SST was measured by directly placing thermometers into buckets of manually collected ocean surface water. This arguably crude method is subject to biases, such as the collected quantity, type of bucket, which are not consistent across all observations. Later, when ship engines became more widely used, the temperature was measured automatically on the water drawn into the engines for cooling. Again, this method has biases, e.g. the ship's structure changing the temperature before it can be measured [142]. Only since the late 70s and early 80s do we have buoys and satellites measuring the SST using Advanced Very High Resolution Radiometer (AVHRR) sensors [22]. Any SST dataset therefore contains data aggregated from a wide range of measurement techniques and differing spatial resolution. The aggregation is performed by correcting for the wide range of measurement techniques, by e.g. estimating biases between measurement methods, interpolating the data into a uniform grid, and applying statistical techniques to ensure consistency across time and space.

3.3 AMV

The following sections will go through the steps for calculating the AMV from gridded SST data. Recall that the AMV are the smoothed, detrended, spatially averaged, SST anomalies in the North Atlantic. The steps for computing the AMV are therefore as follows:

1. Calculate SST anomalies (SSTa).
2. Subset the grid on the North Atlantic basin.
3. Take a spatially weighted average of the SSTa.
4. Detrend with linear regression or remove global mean SST.
5. Smoothing.

3.3.1 Anomalies

The first step is to calculate the anomalies from the SST values. An anomaly is the difference between the absolute value and a baseline value. Within climate science, temperature anomalies are preferred over absolute values because we are interested in the **change** to a reference value instead of the absolute value itself. In addition, statistical summaries of anomalies are not as sensitive in the same way that absolute values are when comparing across locations and seasons [111]. For example, the temperature in Rome in July is warmer than January. Comparing the absolute temperature across these two months would not be meaningful because Julies tend to be warmer than Januaries and the difference would primarily reflect the seasonal difference and not the underlying climate trend. Instead, using temperature anomalies - where the deviation from a long-term average for each month is considered - we would see how the temperatures vary with respect to their respective historical norms. If the anomaly for July is +1.2°C and for January is +1.1°C, it indicates that both months are warmer than their historical averages by similar margins, despite the absolute temperatures being vastly different. This method effectively removes the influence of seasonal variability and helps isolate long-term climatic changes. Furthermore, anomalies enable us to compare temperature changes across different regions with varying climates. For instance, a location in the Arctic may have an absolute winter temperature of -20°C and a summer temperature of 5°C, while a tropical location may range between 25°C and 30°C. Despite these differences, both regions may exhibit similar temperature anomalies over time, revealing patterns that might not be immediately evident from absolute temperature data alone. Temperature anomalies also help in detecting long-term climate trends. A small but consistent upward shift in anomalies over time can indicate global warming, whereas absolute temperature values would be influenced by regional variations, making trends harder to discern. Hence, we need to calculate the anomalies taking into account seasonal variations for each grid point. Let $\mathrm{SST}_{\phi,\theta,y,m}$ be the absolute SST at longitude ϕ, latitude θ, year y and month m. The average seasonal temperature of a month m at grid point ϕ, θ is defined as the weighted sum of all SST values for that month and grid point, for the set of years $y \in Y$ we want to consider.

$$\overline{\mathrm{SST}}_{\phi,\theta,m} = \frac{1}{|Y|} \sum_{y \in Y} \mathrm{SST}_{\phi,\theta,y,m}$$

Essentially, we are calculating the average temperature in January, Feburary, March, etc. for each grid point, over a specific time period. This period is usually 30 years [169] to filter out interannual variations, such as those caused by phenomena like El Niño, while still being short enough to reflect longer-term climatic trends. By subtracting the average seasonal temperature over this baseline period from each grid point for each month, we have then accounted for seasonal variations and can compare values at different grid points with each other.

$$\mathrm{SSTa}_{\phi,\theta,y,m} = \mathrm{SST}_{\phi,\theta,y,m} - \overline{\mathrm{SST}}_{\phi,\theta,m}$$

At a single point in time y, m, we then end up with this matrix $\mathbf{SSTa_{y,m}} \in \mathbb{R}^{n,p}$ of np grid points, where the entry $\mathrm{SSTa}_{y,m,i,j}$ corresponds to the SSTa at longitude i and latitude j in year y and month m.

$$\mathbf{SSTa_{y,m}} = \begin{bmatrix} \text{SSTa}_{1,1,y,m} & \text{SSTa}_{1,2,y,m} & \cdots & \text{SSTa}_{1,p,y,m} \\ \text{SSTa}_{2,1,y,m} & \text{SSTa}_{2,2,y,m} & \cdots & \text{SSTa}_{2,p,y,m} \\ \vdots & \vdots & \ddots & \vdots \\ \text{SSTa}_{n,1,y,m} & \text{SSTa}_{n,2,y,m} & \cdots & \text{SSTa}_{n,p,y,m} \end{bmatrix}$$

3.3.2 Spatial Boundary

The AMV is commonly defined as the spatial average of detrended SSTa in the North Atlantic, however there is no broad consensus for defining the North Atlantic boundaries as well as for detrending the SSTa values. As for the boundaries, Enfield et al. [42] and Trenberth and Shea [153] define them over 0 - 60N and 80W - 0, Sutton and Hodson [144], Palter [119], and Ting et al. [148] use 0 - 60N and 75W - 7.5W, whereas Ruiz-Barradas et al. [130] use 5 - 75W and 0 - 60N (see table 3.2). None of the authors explicitly justify their choice, however most agree on the latitudes and only disagree by 5-7.5 degrees of longitude. In Figure 3.2 we can see that the additional 5-75 degrees in longitude East means to include the North Sea in the calculations (see rectangles 2 and 4), whereas 5 degrees longitude West means we include more of the Carribbean.

Rectangle	Latitude	Longitude	Source
1	0-60N	80W-0E	[42]
2	0-60N	75W-7.5E	[144, 119, 148]
3	0-65N	80W-0E	[151, 153, 115]
4	0-60N	75W-5E	[130]

Table 3.2 The different latitude and longitude boundaries used to defined the North Atlantic.

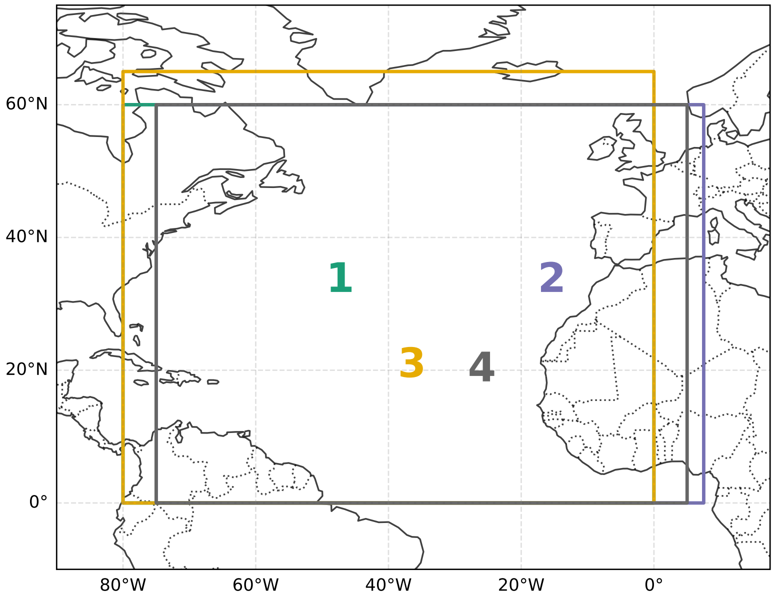


Figure 3.2 Geographical representation of the various boundaries that can be used to define the North Atlantic. Each boundary box is assigned a number that matches its position in the table 3.2.

3.3.3 Spatial Average

From now on, I will use t to index the time dimension instead of the year and month y, m since we only needed to refer to a specific year and month for defining the seasonal detrending. Each entry in the matrix $\mathbf{SSTa_t}$ represents a grid point on Earth and as a next step, we want to get the average value of that matrix, which means taking the **spatial average** at a single point in time. A simple approach would be to just sum all the entries and divide by the number of entries (see equation 3.1), however this would distort the average in favour of higher latitudes.

$$\overline{\text{SSTa}}_t = \sum_{\theta \in \Theta} \sum_{\phi \in \Phi} \mathbf{SSTa}_{t,\theta,\phi} \tag{3.1}$$

Grid points represent different sized areas due to the curvature of the Earth, because the longitude-latitude coordinate system is based on **degrees** and not on **area**. Longitude lines converge to a single point at the poles and become more widely spaced as you move towards the equator. Therefore, a degree of longitude at the poles covers 0km whereas at the equator it is more than 100km. Hence, a simple arithmetic mean would disproportionately weight areas near the poles, where grid cells are smaller and thus denser. To account for this, we compute a **spatially weighted** average over the region 0-60N and 80W-0, where each grid point is weighed by the actual area it represents. Given that the observations are uniformly distributed in latitude, the weights for an observation can be calculated by the cosine of the obversation's latitude, since the area represented by a grid point is proportional to that value [52, 40, 59]. The weight w_θ for observations at latitude θ is equal to $cos(\theta)$ (Figure 3.3).

$$w_\theta = cos(\theta) \tag{3.2}$$

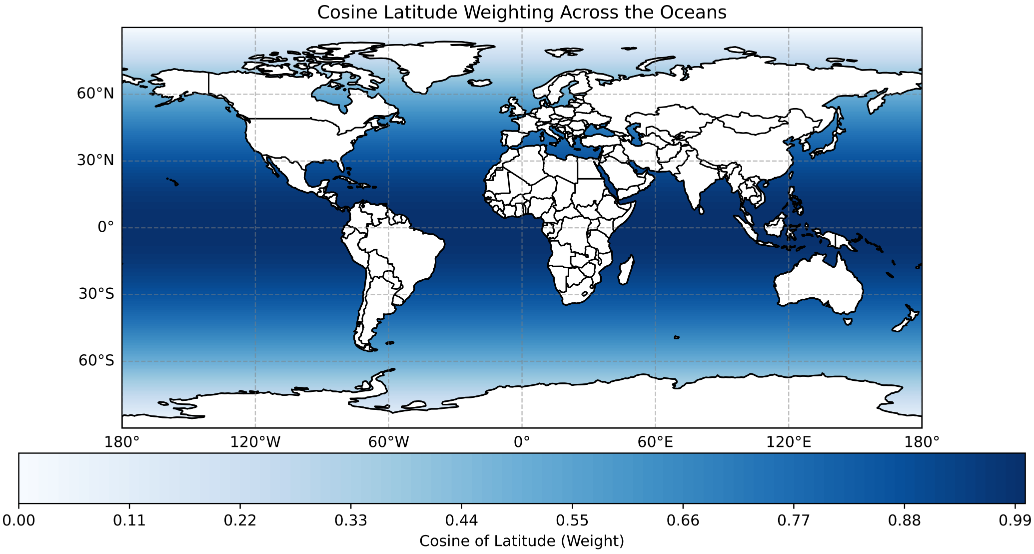


Figure 3.3 The weight assigned to each grid point by the cosine function. Darker areas represent larger weights, lighter areas smaller weights. Near the poles, the weight w_θ is almost 0 to account for the denser observations, whereas at the equator w_θ is 1.

This ensures that the final scalar value accurately reflects the overall quantity across the entire spatial domain without overemphasising regions with finer resolution. The spatial average of the discrete SSTa at the point in time t is then defined as the weighted sum of the SSTa at each grid point weighed by the cosine of its latitude, for all grid points. Θ and Φ represent the set of latitudes and longitudes respectively.

$$\overline{\mathrm{SSTa}}_t = \sum_{\theta \in \Theta} \sum_{\phi \in \Phi} w_\theta \mathbf{SSTa}_{t,\theta,\phi}$$

3.3.4 Detrending

We now have a 1D timeseries $\mathbf{SSTa} = \left[\overline{\mathrm{SSTa}}_1, \overline{\mathrm{SSTa}}_2, \cdots, \overline{\mathrm{SSTa}}_n\right]$ of the spatially averaged SSTa values. The next step is to *detrend* this timeseries to remove the long term warming trend and other external influences such that only the internal variability of the North Atlantic is preserved in the timeseries. Broadly, there are two ways this is usually done. The original approach by Enfield et al. [43] is to fit a linear regression model on the timeseries and retaining the residuals as the detrended timeseries. This approach remains in use, for example Bellomo et al. [4] fit a linear regression, even though it has drawn some criticism which will be presented below. The linear regression is simply the matrix-vector multiplication of the input $\mathbf{X}$ and the transposed weights β^T. Note that for a timeseries $\mathbf{Y} = [y_1, y_2, \ldots, y_t]$, we add a vector of 1 into the input $\mathbf{X} \in \mathbb{R}^{n \times 2}$ to account for the bias term β_1. This simplifies the multiplication with the weight vector $\beta \in \mathbb{R}^2$ containing the bias.

$$\mathbf{X} = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \quad , \quad \beta = [\beta_1, \beta_2] \quad , \quad \mathbf{y} = \mathbf{X}\beta^T + \epsilon$$

The regression function $f(\mathbf{X})$, where $\mathbf{X} \in \mathbb{R}^{n\times 2}, \beta \in \mathbb{R}^{1\times 2}$, is simply the matrix-vector multiplication $\mathbf{X}\beta^T$:

$$f(\mathbf{X}; \beta) = \mathbf{X}\beta^T$$

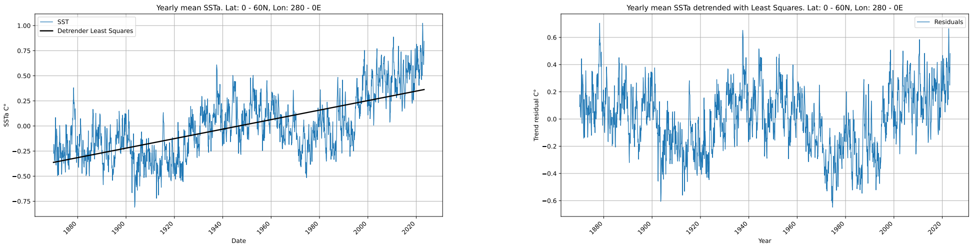
which predicts the values of the timeseries based on the input $\mathbf{X}$, containing the timesteps. Using $f(\mathbf{X})$, we calculate the residuals as the difference between the predicted values and actual observations:

$$\text{residuals}_{\text{regression}} = \mathbf{y} - f(\mathbf{X})$$

Applying this to the SSTa, we obtain:

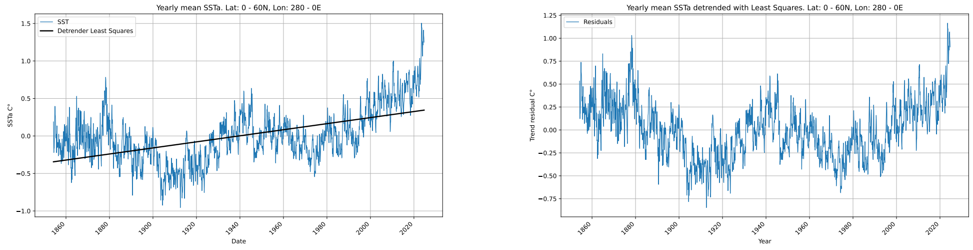
$$\text{SSTa}_{\text{residuals}_{\text{regression}}} = \text{SSTa} - f(\mathbf{X})$$

Figure 3.4 shows fitted regression lines and residuals for the HadISST and ERSSTV5 datasets.



(a) The SSTa and line fitted by the linear regression $f(\mathbf{X})$.

(b) The residuals $\text{SSTa}_{\text{residuals}}$.



(c) The SSTa and line fitted by the linear regression $f(\mathbf{X})$.

(d) The residuals $\text{SSTa}_{\text{residuals}}$.

Figure 3.4 The trend line and regression residuals, calculated from the HadISST 3.4a, 3.4a and ERSSTV5 3.4c,3.4d datasets.

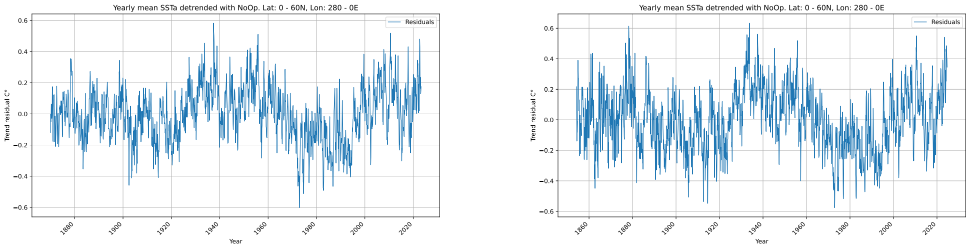
A simpler method is to subtract the global mean SSTa from each grid point, the idea being here that the global mean SSTa acts as a proxy for other external forcing trends [151, 153]. The global mean SSTa is the average SSTa over the spatial dimension. Usually when calculating global averages related to SST, global refers to the extrapolar oceans encompassed by latitudes 60S - 60N [153, 137, 20, 124, 121]. The reason for this being that the polar regions have data gaps due to the changing ice cover, which would contribute noise to the average and distort it. Hence, to calculate the global mean SSTa, we filter the SSTa values by latitude, only keeping values where $\theta \in [-60, 60]$. Given the SSTa values at time t, defined as the matrix $\mathbf{SSTa}_t$, we then calculate a spatial mean over latitudes 60S to 60N and longitudes $[0, 360]$.

$$\overline{\mathrm{SSTa}}_t^{\mathrm{global}} = \sum_{\theta \in [-60,60]} \sum_{\phi \in [0,360]} w_\theta \mathbf{SSTa}_{t,\theta,\phi} \tag{3.3}$$

And subtract this from each entry in the SSTa matrix to get the residuals:

$$\mathrm{SSTa}_{\mathrm{residuals}_{\mathrm{global\text{-}mean}}} = \overline{\mathrm{SSTa}_t} - \overline{\mathrm{SSTa}}_t^{\mathrm{global}} \tag{3.4}$$

Figure 3.5 shows the residuals after the global mean SST was removed, for the HadISST and ERSSTV5 datasets.



(a) The residuals $\mathrm{SSTa}_{\mathrm{residuals}_{\mathrm{global\text{-}mean}}}$ from HadISST.

(b) The residuals $\mathrm{SSTa}_{\mathrm{residuals}_{\mathrm{global\text{-}mean}}}$ from ERSSTV5.

Figure 3.5 The global-mean removal residuals, calculated from the HadISST and ERSSTV5 datasets.

3.3.5 Smoothing

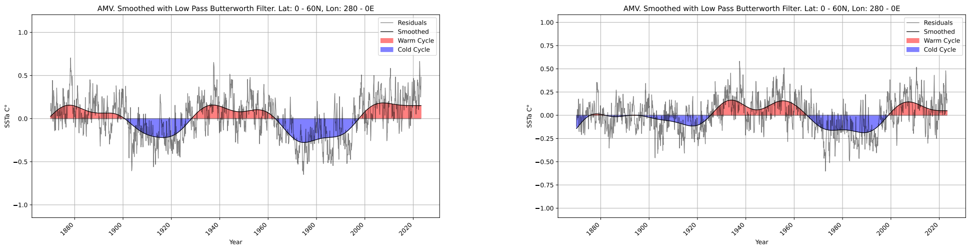
At this stage, the timeseries exhibit short-term fluctuations and high frequency variability which overpower the long-term signals and trends. Smoothing makes these low-frequency trends more visible which helps when dealing with signals that are decadal or multidecadal in scale. For smoothing the raw AMV signal, the most common techniques are:

1. Low pass filter [153, 144, 26, 119, 18]. A low pass filter keeps frequencies **below** a certain cutoff ω_c and **dampens** any frequencies above the cutoff ω_c. The range of frequencies that the filter allows to pass, is called the *passband*, whereas the frequencies that the filter weakens is called the *stopband*. The filter works by applying a Fast Fourier Transform (FFT) to the timeseries to get the signal's amplitude at different frequencies, filtering the frequencies, and converting the signal back into the time domain. A common choice is the

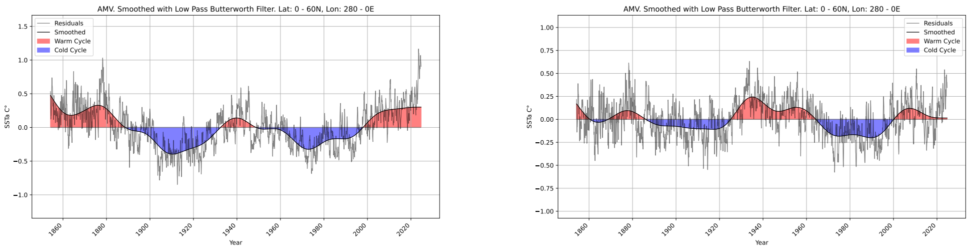
Butterworth filter, which provide a smooth frequency response, meaning that the transition from passband to stopband is not sudden but occurs gradually and continuously. The magnitude response of the filter $|H(j\omega)|$ tells us how much the filter affects frequency component strength of the input.

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2N}}} \tag{3.5}$$

Figure 3.6 shows the Butterworth filter applied to the raw AMV index calculated from the ERSSTV5 and HadISST datasets using linear detrending and global-mean removal.



(a) HadISST with linear detrending. **(b)** HadISST with global mean removal.



(c) ERSSTV5 with linear detrending. **(d)** ERSSTV5 with global mean removal.

Figure 3.6 The smoothed AMV signal, computed from HadISST and ERSSTV5 datasets using linear detrending and global mean removal. The black line is the Butterworth filtered raw AMV index (grey line). Red areas show warm cycles, blue areas show cold cycles of the index.

2. 10 Year Running Mean [19, 150, 1, 12, 43] (Knudsen [76] applies a 5 year running mean but, 10 years is more frequently used). The running mean is computed as a series of averages, where each value is obtained from a **sliding window** of size k applied to the time series (Figure 3.7). Mathematically, this is a **1D convolution** operation, where we convolve the time series $\mathbf{y} = (y_1, \ldots, y_t)$ with a uniform kernel

$$\mathbf{h} = \left[\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}\right] \tag{3.6}$$

of size k. The running mean at time step i, denoted as x_i, is computed as the **dot product** between the kernel **h** and the corresponding window of the time series. Specifically, at index i, we take the segment

$$\mathbf{y}_{i-\lfloor \frac{k}{2} \rfloor : i+\lfloor \frac{i}{2} \rfloor} \tag{3.7}$$

and compute

$$x_i = \mathbf{h}^T \mathbf{y}_{i-\lfloor k/2 \rfloor : i+\lfloor k/2 \rfloor}. \tag{3.8}$$

In essence, since each entry in the kernel is equal to $\frac{1}{k}$, by doing a dot product we are taking an arithmetic mean. This ensures that each value in the smoothed series is an average over a symmetric window around i. A drawback with this approach is that the smoothed timeseries is shorter than the original one, since we cannot compute a mean for the first and last $\lfloor \frac{k}{2} \rfloor$ entries, since the kernel would ”protrude” over the timeseries. This can be solved by padding the edges with constants or some statistical summary of the near-edge data, however these methods all introduce some biases.

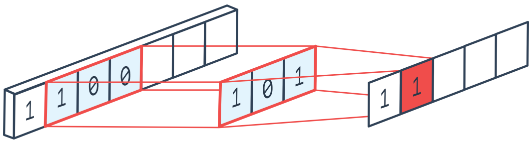
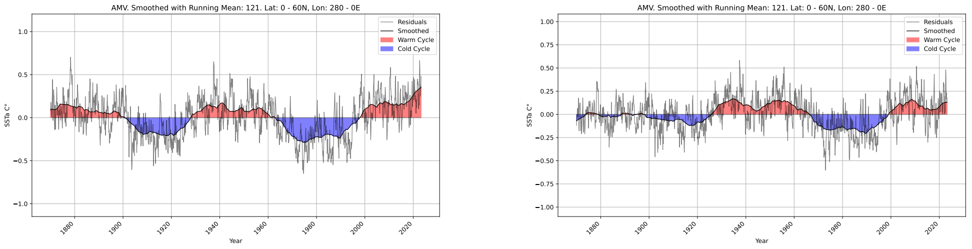
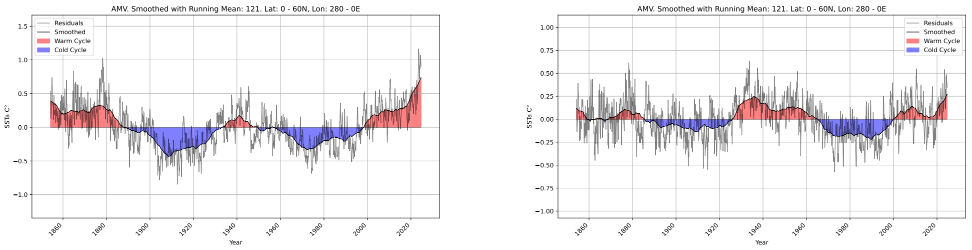


Figure 3.7 Example of a 1D convolution with $k = 3$. Image source

Figure 3.8 shows the running-mean smoothed AMV signal computed from HadISST and ERSSTV5 datasets using linear detrending and global mean removal.



(a) HadISST with linear detrending. **(b)** HadISST with global mean removal.



(c) ERSSTV5 with linear detrending. **(d)** ERSSTV5 with global mean removal.

Figure 3.8 The smoothed AMV signal, computed from HadISST and ERSSTV5 datasets using linear detrending and global mean removal. The black line is the 121 month running mean of the raw AMV index (grey line). Red areas show warm cycles, blue areas show cold cycles of the index.

3.3.6 NOAA Comparison

Finally, we compare different methods for computing the AMV index, including linear detrending versus global mean removal, smoothed versus raw data, and HadISST versus ERSSTv5, against the AMV index provided by NOAA. NOAA provides a smooth and raw version of the AMV index, available online, that we compare against the calculated indices. The following sections present plots of the various calculated indices alongside NOAA's AMV index, as well as tables summarizing error metrics. To assess the accuracy of the computed AMV index, we use Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The error functions are defined as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |f(t) - y_t| \tag{3.9}$$

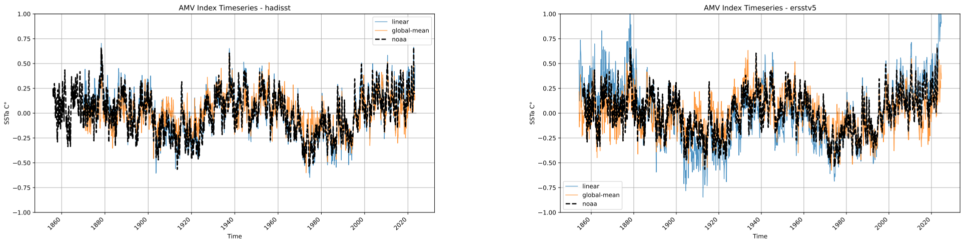
$$MSE = \frac{1}{N}\sum_{t=1}^{N} (f(t) - y_t)^2 \tag{3.10}$$

$$RMSE = \sqrt{MSE} \tag{3.11}$$

where $f(t)$ is the computed index and y_t the precomputed index by NOAA.

3.3.7 Raw AMV Index

The raw index is the seasonally detrended, spatially averaged, detrended SSTa temperature in the North Atlantic, without any smoothing applied. Figure 3.9b and 3.9b show the raw index calculated from the HadISST and ERSSTV5 datasets respectively, against the raw index by NOAA. The plots are difficult to judge since the raw index is quite noisy, the table 3.3 however shows that the method and dataset with the lowest error for all three functions is **linear detrending** on the **HadISST** dataset, being almost half of the other errors.



(a) HadISST **(b)** ERSSTV5

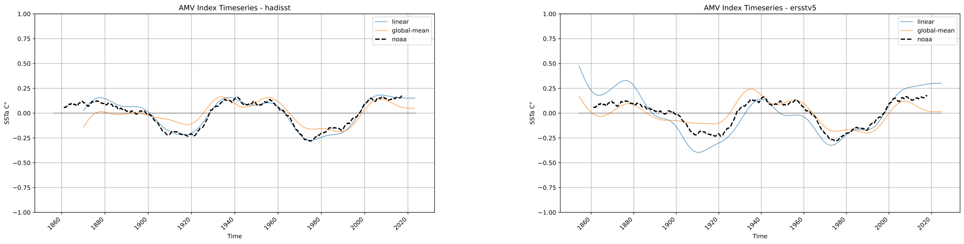
Figure 3.9 The raw AMV signal, calculated from the HadISST and ERSSTV5 datasets, with linear detrending (blue line) and global mean removal (orange line), compared against the raw AMV index by NOAA (black line).

Dataset	Method	MAE	MSE	RMSE
hadisst	global-mean	0.105622	0.017876	0.133702
hadisst	linear	**0.060146**	**0.006022**	**0.077599**
ersstv5	global-mean	0.117449	0.022020	0.148393
ersstv5	linear	0.120107	0.023899	0.154594

Table 3.3 Comparison of raw AMV index calculations to raw AMV index by NOAA, using error functions MAE, MSE, and RMSE. Bold values indicate the lowest value in each column.

3.3.8 Smoothed AMV Index

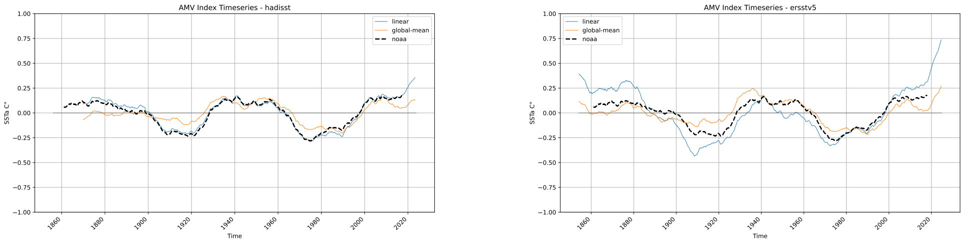
For comparing the smoothed indices, we have an additional feature to compare, namely the smoothing technique. Figures 3.10a and 3.10b show the AMV index smoothed with the Butterworth low-pass filter, figures 3.11a and 3.11b show the AMV index smoothed with a 121-month running mean. The table 3.4 shows the error for all the different combinations of dataset (HadISST or ERSSTV5), method (linear detrending or global mean removal) and smoothing technique (low-pass filter or running mean). Again, the combination with the smallest error is the linear detrending on the HadISST dataset, smoothed with the 121 month running mean.



(a) HadISST

(b) ERSSTV5

Figure 3.10 The low-pass smoothed HadISST and ERSSTV5 signals, with linear detrending (blue line) and global mean removal (orange line), compared against the smoothed AMV index by NOAA (black line).



(a) HadISST

(b) ERSSTV5

Figure 3.11 The 121 month running mean smoothed HadISST and ERSSTV5 signals, with linear detrending (blue line) and global mean removal (orange line), compared against the smoothed AMV index by NOAA (black line).

Dataset	Method	MAE	MSE	RMSE	Smoothing
hadisst	global-mean	0.064478	0.005993	0.077417	low-pass
hadisst	linear	0.025511	0.000971	0.031158	low-pass
ersstv5	global-mean	0.065464	0.005881	0.076688	low-pass
ersstv5	linear	0.091259	0.011821	0.108723	low-pass
hadisst	global-mean	0.063881	0.005848	0.076473	running-mean
hadisst	linear	**0.020614**	**0.000649**	**0.025473**	running-mean
ersstv5	global-mean	0.064634	0.005763	0.075914	running-mean
ersstv5	linear	0.089849	0.011356	0.106564	running-mean

Table 3.4 Comparison of smooth AMV index calculations to the smooth AMV index by NOAA. Bold values indicate the lowest value in the column.

3.4 PDO

The steps for calculating the PDO are similar to the AMV calculation:

1. Remove global mean SST from each grid point.

2. Calculate SST anomalies (SSTa).
3. Subset the grid on the North Pacific basin.
4. Perform EOF/PCA and select first principal component.
5. Standardise the first principal component.
6. Smooth using running mean.

3.4.1 Remove Global Mean

As a first step, we remove the global mean SST, calculated over $60S - 60N$, from each grid point. The concept is the same as the global mean SSTa removal for the AMV (see section 3.3.4), however instead of subtracting the global mean SSTa, we subtract the global mean SST at time t from each grid point at time t. Shown below, the global mean SST, $\overline{\mathrm{SST}}_t^{\mathrm{global}} \in \mathbb{R}$ is the average SST over latitudes $60S - 60N$ and longitudes $0 - 360$, where each grid point is weighed by its area weight w_θ (see section 3.3.3).

$$\overline{\mathrm{SST}}_t^{\mathrm{global}} = \sum_{\theta \in [-60,60]} \sum_{\phi \in [0,360]} w_\theta \mathbf{SST}_{t,\theta,\phi} \tag{3.12}$$

We then subtract this from each grid point, doing an element wise subtraction from the matrix $\mathrm{SST}_t \in \mathbb{R}^{\theta \times \phi}$.

$$\mathbf{SST}_t - \overline{\mathrm{SST}}_t^{\mathrm{global}} \tag{3.13}$$

3.4.2 Anomalies

This step is the same as for the AMV (see section 3.3.1), we subtract the climatological mean (calculated over a reference period) from each month, e.g. subtract the average January SST in the period 1971-2000 from each January in the dataset.

3.4.3 Spatial Boundaries

Like with the AMV, there are different ways to define the boundaries of the North Pacific basin. Table 3.5 and Figure 3.12 show the various boundaries found in the literature. In our experiments, we define the Pacific domain using longitude bounds from 110E to 100W, following previous studies [152, 125, 173]. This range encompasses most of the northern South China Sea and extends to the western coast of Mexico. For latitude, we use 20N to 65N. While 20N is a consistent southern boundary across the literature, the northern boundary varies, with some studies extending beyond the Bering Strait. Here, we choose to place the northern boundary precisely at the Bering Strait as in [23], which serves as a natural division between the Pacific and Arctic Oceans.

Rectangle	Latitude	Longitude	Source
1	20 - 70N	110E - 100W	[152, 125, 173]
-	20 - 70N	?	[27, 108]
2	20 - 60N	120E - 110W	[164]
3	20 – 65N	120E - 100W	[23]

Table 3.5 The various latitude and longitude boundaries used to define the North Pacific.

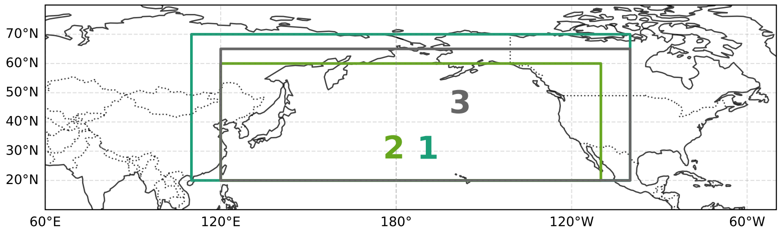


Figure 3.12 The boundary boxes from table 3.5 plotted on a map. The numbers correspond to the rows in 3.5.

3.4.4 EOF/PCA

After selecting the North Pacific basin, we will compute EOF/PCA on the SSTa data. From here on, I will simply refer to this step as PCA. PCA is a dimensionality reduction technique that projects the data onto a lower-dimensional space by doing an eigenvalue decomposition and then projecting the data onto the obtained eigenvectors. The projection onto the eigenvectors gives us the principal components (new features), projecting onto the first eigenvector gives the first principal component, projecting onto the second gives the second principal component etc. The principal components are ordered by the amount of variance they contribute, from most to least. Since the eigenvectors form an orthornomal basis, the principal components obtained are also orthogonal to each other.

Recall that our SSTa data at this stage is in $\mathbb{R}^{t \times \theta \times \phi}$, meaning that we cannot directly apply PCA to it as is since the eigendecomposition is designed to work with 2D matrices and not higher dimensional tensors. The solution is to **collapse** the latitude and longitude dimensions into a single dimension such that the data is in $\mathbb{R}^{t \times n}$ where $n = \theta\phi$ [55], and each timestep is now a **vector** in $\mathbb{R}^n$. Given that $n = \theta\phi$, the function that does the concatenation can be expressed as $f : \mathbb{R}^{t \times \theta \times \phi} \to \mathbb{R}^{t \times n}$. The final data matrix $\mathbf{X}$ then looks as follows:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t1} & x_{t2} & \cdots & x_{tn} \end{bmatrix}$$

Each column in $\mathbf{X}$ now corresponds to a point in space and tells us how that point varies through time. Using Singular Value Decomposition, we can write the data matrix $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{U}\,\mathbf{V}^T \tag{3.14}$$

Where, starting from the right:

- $\mathbf{V}^T \in \mathbb{R}^{n \times n}$ has orthornomal columns that are the **right singular vectors**, which are eigenvectors representing the **spatial patterns** found in the data. The eigenvector is in an n-dimensional space where each axis is a grid point, hence the direction of the vector tells us .... The entries in the vector tell us how the spatial patterns vary, e.g. if all entries in the vector have the **same sign**, the grid points vary **together**. The magnitude of the entries tell us how each grid point contributes to this pattern.
- $\Sigma \in \mathbb{R}^{t \times n}$ is a diagonal matrix of **eigenvalues**, which tell us how much variance is explained by each pattern in $\mathbf{V}^T$.
- $\mathbf{U} \in \mathbb{R}^{t \times t}$ contain the **left singular vectors** which represent how the spatial patterns in $\mathbf{V}^T$ evolve over time, how strong or weak the pattern is at each timestep.

By applying PCA and selecting a subset of the axes that make up the orthonormal space, we are essentially reducing the spatial dimensions to axes that account for the most variability. where the principal components show us the strength of these patterns. Figure 3.13 illustrates the steps for calculating the PDO from the SSTa data.

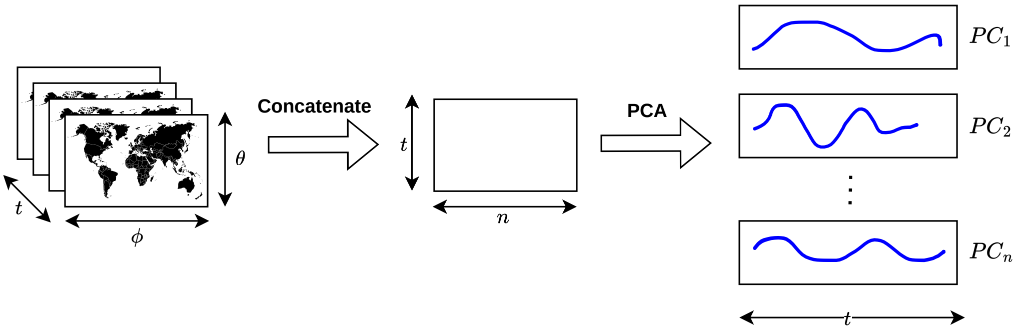


Figure 3.13 Starting from the data in $\mathbb{R}^{t \times \phi \times \theta}$, we first concatenate the longitude and latitude dimensions into a single dimension of length n, resulting in a matrix in $\mathbb{R}^{t \times n}$. By applying the PCA algorithm to this data we end up with n principal components, each of which is a timeseries of length t. The first principal component represents the PDO, hence we choose this timeseries for the following analyses.

3.4.5 Standardise

Now that we have the first principal component, we **standardise** it to $\mu = 0$, $\sigma = 1$. Let $PC_1 \in \mathbb{R}^t$ be the first principal component obtained by PCA, then we do the following:

$$\frac{PC_1 - PC_1^{\mu}}{PC_1^{\sigma}} \tag{3.15}$$

where PC_1^{μ} is the mean of the principal component timeseries and PC_1^{σ} is the standard deviation, defined below, t is the length of the timeseries.

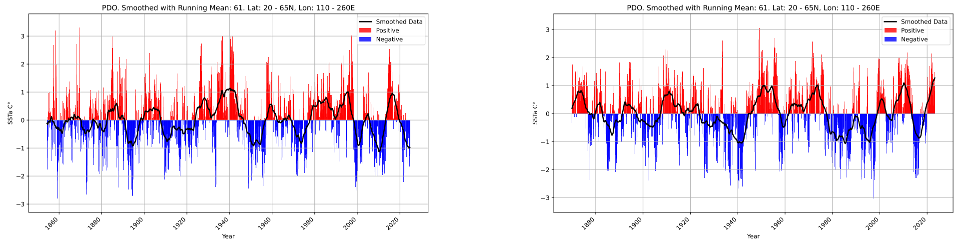
$$PC_1^{\mu} = \frac{1}{t}\sum_{i=1}^{t} PC_{1,i} \tag{3.16}$$

$$PC_1^{\sigma} = \sqrt{\frac{1}{t}\sum_{i=1}^{t} (PC_{1,i} - PC_1^{\mu})^2} \tag{3.17}$$

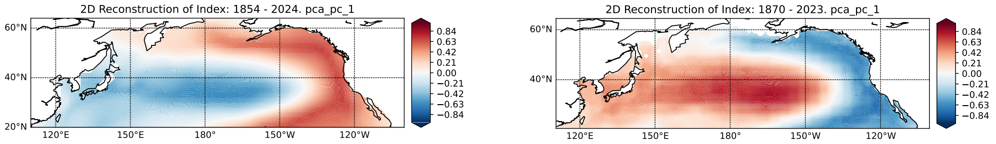
3.4.6 Smoothing

Finally, just like the AMV, we apply smoothing to the resulting time series. The standard approach in the literature is a running mean (see Section 3.3.5). However, because PDO cycles are shorter than those of the AMV - ranging from annual to decadal timescales - the chosen window size for smoothing is also smaller. A common choice is approximately half the window size used for the AMV (121 months). Trenberth and Fasullo, for instance, use a 61-month window [152], while Mantua and Hare, Deser et al., and the Japan Meteorological Agency (JMA) favor a 5-year window (60 months) [92, 27, 65]. Some studies opt for even shorter windows; Yao et al. use a 6-month window [174], while Zhang et al. apply a 9-month running mean followed by a 5-month running mean [181]. We choose to use a 61 month window, since it aligns with the 5 year windows by [92, 27, 65] ($5 \times 12 = 60$) when considering that to centre the window on a timestep $PC_{1,i}$, we require an odd window size to ensure that we consider the same number of past and future values.

Fig 3.14a and 3.14b show the PDO - the standardised first principal component in the region 20-60N, 110E - 100W - calculated from the ERSSTV5 and HadISST datasets respectively. Red indicates when the index is in its warm phase, blue when the index is in its cold phase. The black line is a 61 month running mean. Figures 3.14c and 3.14d show the regression coefficients β of a linear regression model $PDO = \beta\mathbf{SSTa} + \epsilon$, where we treat the PDO index timeseries as a target Y and the individual grid point SSTa's as the covariates or independent variable X. This is a common way to show the spatial pattern of the PDO. Areas with red values are strongly linked to the phase of the PDO, they move together. Blue areas on the other hand correlate negatively, and therefore show SSTa regions that are in opposite phase to the PDO, how they cool when the PDO is in a warm phase and vice versa.



(a) ERSSTV5 **(b)** HadISST

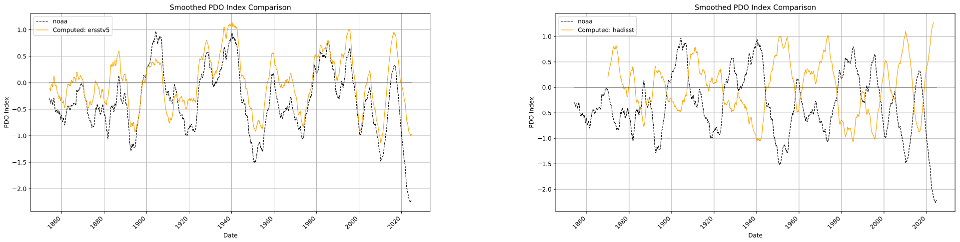


(c) ERSSTV5 **(d)** HadISST

Figure 3.14 The PDO index calculated from the ERSSTV5 and HadISST datasets and the correlation maps. The spatial map for HadISST is inversed because the index computed from the HadISST dataset is basically the inverse of the NOAA index used for comparison, see Figure 3.15b.

3.4.7 NOAA Comparison

As for the AMV, we compare the computed indices from the ERSSTV5 and HadISST datasets with the PDO provided by NOAA. The PDO index computed from the ERSSTV5 dataset (Figure 3.15a) has a lower error than the one calculated from the HadISST dataset. In fact, the HadISST PDO index appears to almost be the inverse of the index provided by NOAA (Figure 3.15b), though the processing pipeline applied is the exact same as for ERSSTV5.



(a) ERSSTV5 **(b)** HadISST

Figure 3.15 Comparison between the computed PDO index from ERSSTV5 and HadISST (orange line) and the PDO by NOAA (dashed black line).

3.5 Subdividing the Pacific

A system consisting solely of the AMV and PDO indices is likely too simplistic to accurately capture the underlying causal relationships. Therefore, it could be useful to include additional variables in the system for several reasons. First, other regions of the Pacific Ocean may exert influence on both the AMV and PDO simultaneously.

This type of interaction, known as a *common cause*, can create the false appearance of a direct causal link between AMV and PDO even when no direct relationship exists. Without accounting for such confounding influences, causal discovery methods may incorrectly infer a connection. Second, causal relationships between AMV and PDO may not be direct, but instead mediated through intermediate processes or regions, known as *pathways*. In this case, AMV could influence another variable, which in turn affects the PDO. Ignoring such mediators could result in missing the true pathway of interaction. Moreover, previous studies have suggested that the AMV can affect variability in the tropical Pacific [149, 66], which is a region that is not fully captured by the PDO index. Incorporating key regions of the tropical and subtropical Pacific into the system could therefore help disentangle direct effects from indirect or spurious ones, leading to a more accurate reconstruction of the causal structure.

There are many ways one could divide the Pacific further. Geographically, it could be divided into equatorial, tropical, and subtropical regions, from which we calculate a spatial average SST, though these might not represent any actual physical phenomena since Pacific indices like ENSO or the PDO are not calculated along latitude lines but SST variability. Another way could be to apply PCA to the entire Pacific basin and pick the first k components as variables. However the resulting principal components of PCA are **orthogonal** to each other, meaning that they are also **uncorrelated**. While this is convenient in a mathematical way, physical processes in the real world are rarely uncorrelated due to the complex nature of the climate system [113]. Since PCA focuses on maximising the variance, it will most likely **mix** underlying physical processes to create the principal components, where at the end each principal component **cannot** be clearly traced back to a physical mode. Additionally, PCA tends to capture global patterns that explain the maximum variance across the entire dataset, whereas physical modes are more localised and do not represent a dominant global variance.

This approach is nonetheless interesting, since it is entirely data driven and does not require any prior assumptions about the spatial structure of the variability. The solution to make PCA better capture local patterns that can be more easily traced back to physical modes, is to rotate the principal components such that their loadings, the weights that represent how much each original variable contributes to the new principal components, are redistributed in a way that makes the principal components more interpretable. This rotation is valled *varimax*. After finding the principal components using PCA from the data matrix $\mathbb{R}^{t \times n}$, where t is the number of time points and n the number of variables, we obtain a set of p principal components. These are associated with a loading matrix in $\mathbb{R}^{n \times p}$, where each entry represents the contribution (or weight) of an original variable to a given component. Each principal component (PC) assigns a loading (or weight) to every original variable. The goal of varimax rotation is to maximize the total variance of the squared loadings across all components. This encourages a structure where each PC has a few variables with large loadings and most with small or near-zero loadings, resulting in a sparse and localised representation. This sparsity often reveals more interpretable and correlated spatial patterns in the data [54, 134, 159, 165]. Varimax-rotated PCA has been successfully applied to identify SST-based predictors of rainfall in Pakistan [116]. The method was originally introduced by

[68] and further developed in the climate context by [158]. Mathematically, the varimax criterion seeks the rotation that maximizes the variance of the squared loadings, which is illustrated in Figure 3.16.

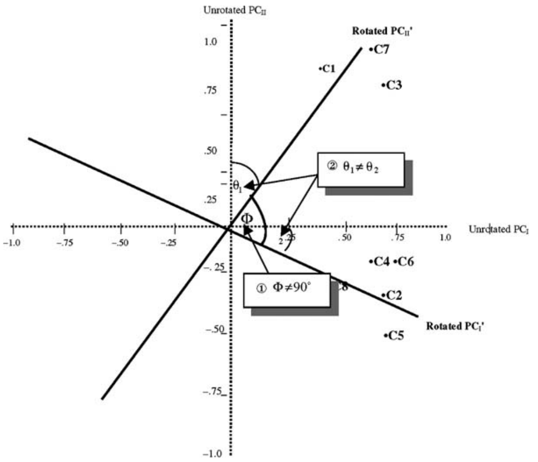
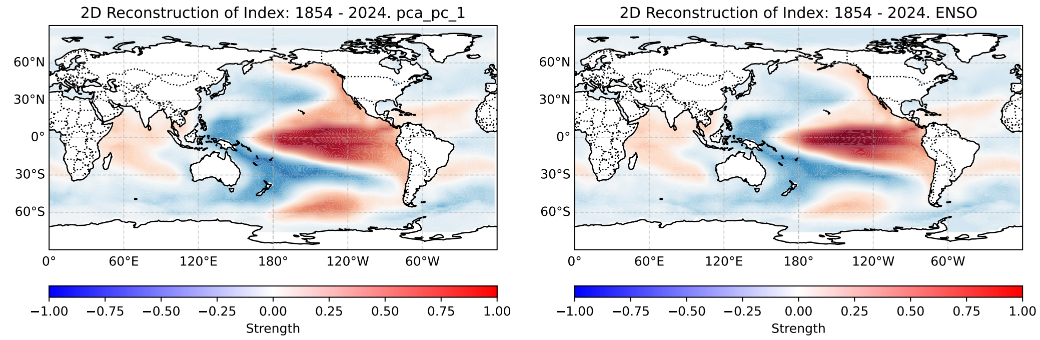


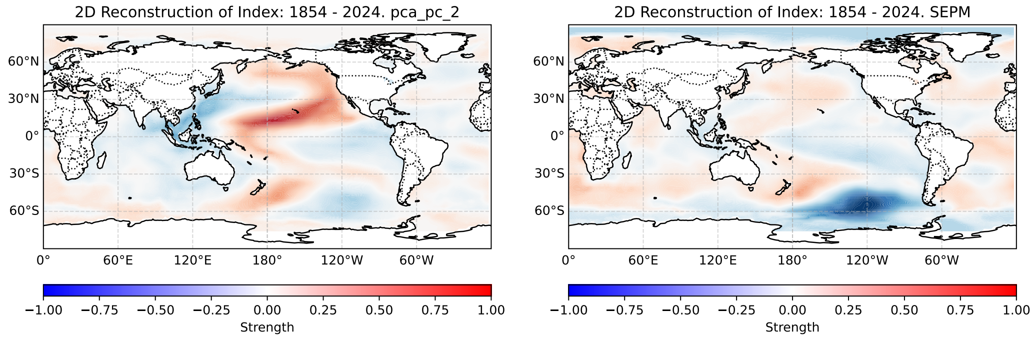
Figure 3.16 Original components (dotted line) and rotated components (solid line)[123].

3.5.1 3 Varimax Rotated Principal Components

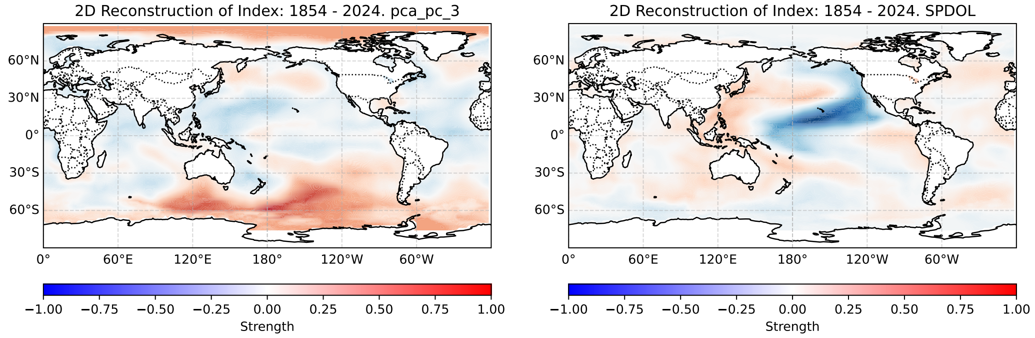
Comparing the spatial correlation maps of the first three principal components in the Pacific with those of the same components after applying a varimax rotation reveals notable differences, particularly for the second and third components. The first component remains virtually unchanged and closely resembles the canonical ENSO pattern, characterized by a tongue of anomalies extending from the west coast of South America into the central Pacific (see Figures 3.17a and 3.17b). Although the rotated version shows slightly more extreme correlation values, the similarity between the two suggests that the loadings of the first principal component were already quite sparse, indicating that this mode corresponds to a well-defined physical phenomenon—namely, ENSO. In contrast, the second and third components undergo substantial changes following the varimax rotation. For the second component, the rotated version exhibits a spatially concentrated pattern in the South Pacific (Figure 3.17d), hence we decided to name it *Southeast Pacific Mode* (SEPM), whereas the unrotated component displays stronger activity in the subtropical North Pacific near the Tropic of Cancer (Figure 3.17c). The third rotated component shows a horseshoe-shaped pattern in the subtropical North Pacific (Figure 3.17f), similar to the second unrotated component. Due to its proximity to the PDO, we named it *Southward PDO Leg* (SPDOL). By comparison, the unrotated third component is weaker and spread across both polar regions (Figure 3.17e), with little spatial concentration.



(a) PC1 **(b)** PC1 Varimax Rotation (ENSO)



(c) PC2 **(d)** PC2 Varimax Rotation (SEPM)



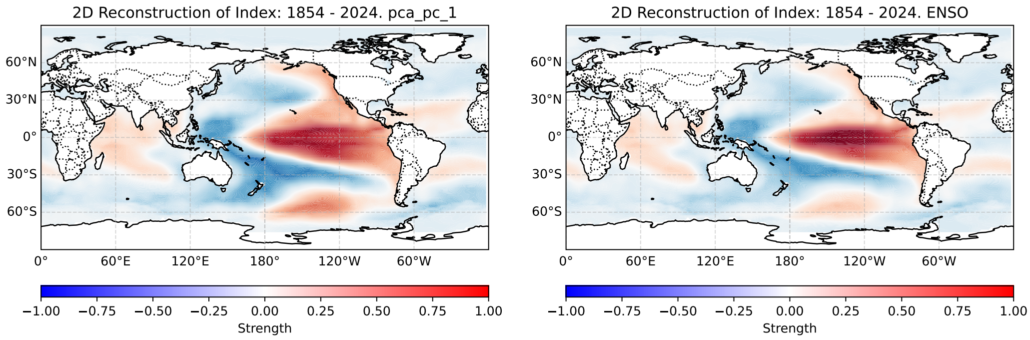
(e) PC3 **(f)** PC3 Varimax Rotation (SPDOL)

Figure 3.17 The spatial maps of the first three principal components obtained from PCA and the first three **varimax rotated** principal components, calculated over the region 60S-60N, 110E - 100W using the ERSSTV5 dataset. The value at each grid point is the Pearson Correlation Coefficient between the SSTa timeseries at that grid point and the principal component.

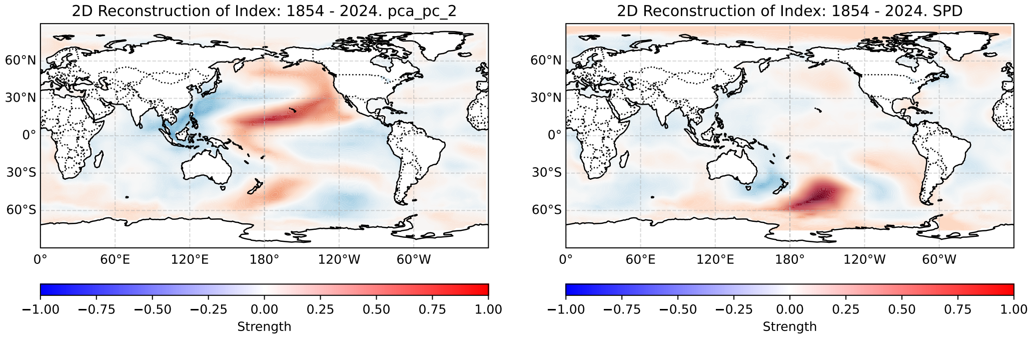
3.5.2 5 Varimax Rotated Principal Components

For the system with five principal components, the varimax-rotated results again reveal the characteristic ENSO pattern in the first component (Figures 3.18a and

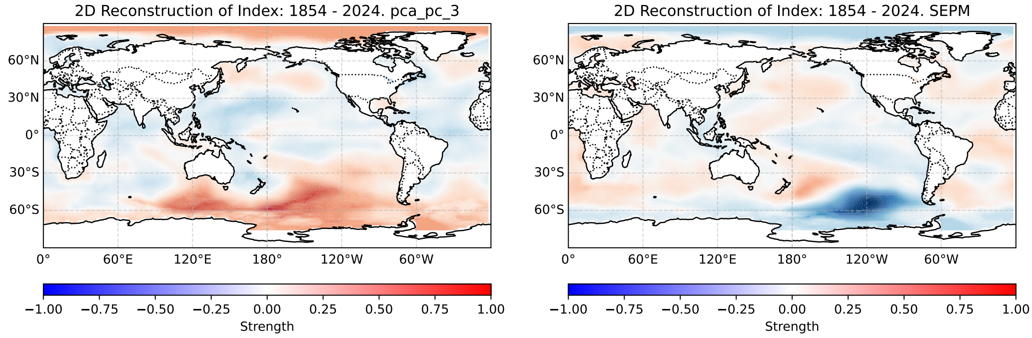
3.18b), consistent with what was observed in the three-component analysis. Notably, the first three unrotated principal components remain essentially unchanged regardless of whether three or five components are retained, as the leading modes of variability captured by PCA are invariant to the number of components selected. However, the varimax-rotated components—except for the first—differ between the three- and five-component systems. This is because varimax rotation operates on the full set of selected components, optimizing the variance of the loadings within the context of that specific system. As a result, increasing the number of components alters the rotational solution. In the five-component system, the second varimax-rotated component remains focused in the South Pacific, but its spatial pattern shifts westward and exhibits an inverted sign in its correlation map compared to its counterpart in the three-component system. Therefore, we call it *South Pacific Mode* (SPM). The third varimax-rotated component corresponds closely to the second rotated component in the three-component system (Figures 3.18f and 3.17d) and therefore is the SEPM. The fourth varimax-rotated component emerges as a well-defined structure southwest of Australia (Figure 3.18h) and is again more spatially concentrated and of greater magnitude than its unrotated equivalent (Figure 3.18g). Due to the proximity to the Tasman Sea, we call it the *Tasman Sea Dipole* (TSD). Lastly, the fifth varimax-rotated component (Figure 3.18j) resembles the third rotated component in the three-component system (Figure 3.17f), so we also call it SPDOL, while the corresponding unrotated fifth component remains diffuse and weaker in amplitude.



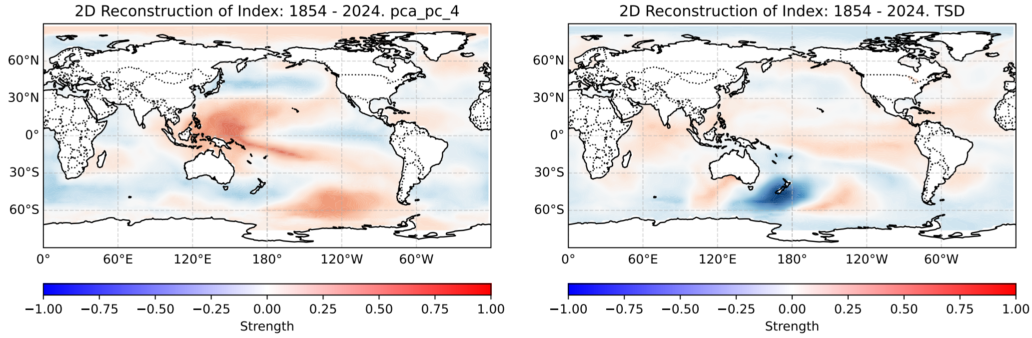
(a) PC1 **(b)** PC1 Varimax Rotation (ENSO)



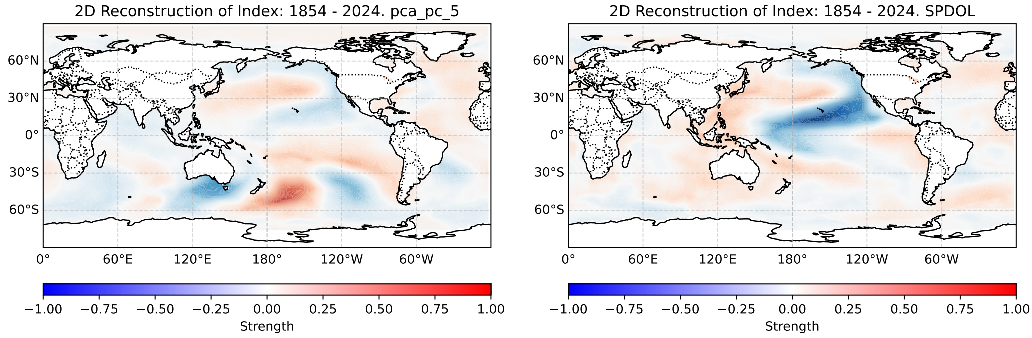
(c) PC2 **(d)** PC2 Varimax Rotation (SPM)



(e) PC3 **(f)** PC3 Varimax Rotation (SEPM)



(g) PC4 **(h)** PC4 Varimax Rotation (TSD)



(i) PC5 **(j)** PC5 Varimax Rotation (SPDOL)

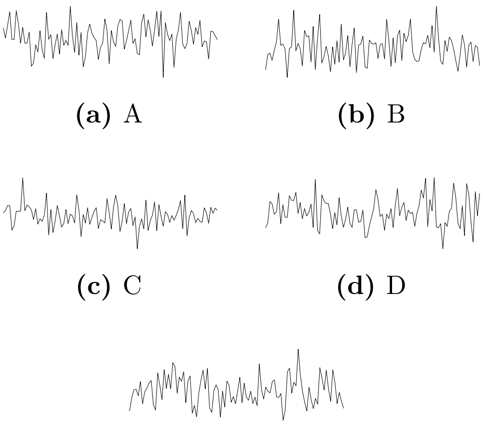
Figure 3.18 The spatial maps of the first five principal components obtained from PCA and the first five **varimax rotated** principal components, calculated over the region 60S-60N, 110E - 100W using the ERSSTV5 dataset. The value at each grid point is the Pearson Correlation Coefficient between the SSTa timeseries at that grid point and the principal component.

Chapter 4

Methods

4.1 Causal Discovery

Causal Discovery methods aim to **discover** and **quantify** relationships between variables, i.e. **is there** a relationship and **how strong** is it? The goal is to infer the causal structure from the data, often represented as a graph. Figure 4.3 illustrates the causal discovery problem - we observe timeseries and want to retrieve causal relationships between the different time series (Figure 4.2). Using causal methods, we can go beyond simple correlation analyses of the system and find as many true causal relationships as possible, all while also minimising the number of spurious links.



(a) A

(b) B

(c) C

(d) D

(e) E

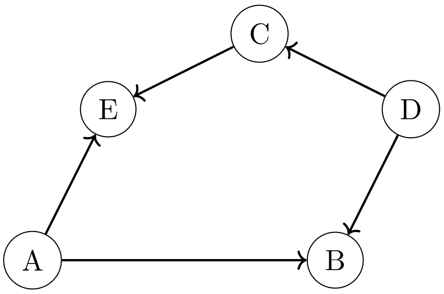


Figure 4.2 $A = f_A()$
$B = f_B(A, D)$
$C = f_C(D)$
$D = f_D()$
$E = f_E(A, C)$

Figure 4.3 Example timeseries and a corresponding causal graph.

4.2 Granger Causality

Out of the three methods in this work, Granger causality was the first method that proposed a quantifiable way of defining and measuring causality for time series data, and is also the method that has seen application in the most diverse range of areas.

Initially applied to the field of econometrics [53], Granger Causality is used in both financial and economic areas [39, 60, 28], and has also found its way into politics [49], and neuroscience [14, 138, 31]. It also remains a popular analysis within earth science [122, 95, 100, 77, 41]. Granger causality has been used to study the AMO and PDO [154, 143]. We are using the `grangercausalitytests` module from the `statsmodels` [136] Python package to implement Granger Causality tests.

Granger causality checks whether a timeseries X can better forecast timeseries Y, than just Y itself [120]. Given two timeseries, $X = [X_1, X_2, ...X_n]$, $Y = [Y_1, Y_2, ...Y_n]$, we fit two linear autoregressive models. The first one (reduced), f_{red} is purely **autoregressive**, meaning that it predicts the value Y_t using only past values of Y up until a certain time lag p. The second one (full), f_{full}, predicts Y_t with the past of Y and the past of X up until time lag q.

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t$$

$$f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t$$

Explanation

f_{red} calculates a weighted sum for the p preceding timesteps of Y, where a_0 is the intercept, a_i are the coefficients for the timesteps and ϵ_t is the error. f_{full} sums the past of Y (coeffiencts b_i) with a term which is the weighted sum of the q preceding timestamps of X where c_j are coefficients. η represents the error of the second model.

We then compare the predictions errors of both models, using a scoring function such as Residual Sum of Squared (RSS).

$$RSS = \sum_{i=1}^{n} (\hat{y} - y)^2 \tag{4.1}$$

where $\hat{y}$ is the prediction and y is the ground truth. We say that X does *Granger cause* Y if the RSS of the full model f_{full} is **less** than the RSS of reduced model f_{red} and this lower error is **statistically significant** at the chosen level α, indicating that the prediction improvement introduced by including the past of the variable Y is not purely random.

$$X \text{ Granger causes } Y \text{ if } RSS_{full} < RSS_{red} \wedge P(F) < \alpha \tag{4.2}$$

The suggested statistical test to use is an F-test [139] that uses the RSS of both models to compute an F-statistic, which is defined as:

$$F = \frac{RSS_{red} - RSS_{full}/(r-s)}{RSS_{full}/(T-r)} \tag{4.3}$$

Explanation

r is the number of parameters in the full model f_{full}, s is the number of parameters in the reduced model f_{red}, T are number of observations in the timeseries. Essentially, this measure quantifies how much better a more complex model f_{full} performs in predicting the data compared to a simpler model f_{red}. It does this by evaluating the reduction in prediction errors when using the more complex model, relative to the errors that still remain even with the more complex model, adjusted for model complexity.

We calculate the probability of observing the F-statistic under the assumption that the full model does **not** improve the prediction, over the reduced model. If this probability is less than the significance level α, we reject the null hypothesis that the full model does not lead to an improved prediction over the reduced model, illustrated in Figure 4.4. This rejection implies that the past values of X do indeed provide valuable information for predicting Y, suggesting that X Granger-causes Y. Essentially, if including the past of X does not improve the prediction of Y at a statistical significance level α, then all the necessary information needed for predicting Y is already contained in the past of Y itself. We conclude that X Granger causes Y if the probability of observing F less than the chosen significance threshold α.

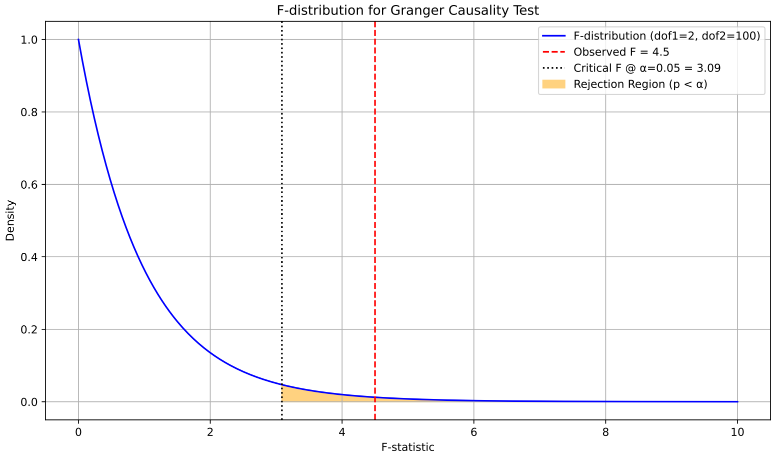


Figure 4.4 Example of an F-distribution with $r = 2, s = 100$, significance level $\alpha = 0.05$ and the observed F-statistic $F = 4.5$. The rejection region (orange area) where the null hypothesis, i.e. that the full model does **not** explain significantly more variance than the reduced model. The vertical dashed line marks the **observed** F-value, which lies within the rejection region, indicating statistical significance at the 5% level

There are several issues with this version of Granger Causality when applying it to earth science data. Since it is a standard linear regression technique, it suffers from the curse of dimensionality for high dimensional datasets, which are common in the field. Though this problem can be addressed with non-linear version of Granger Causality like Neural Networks [146]. This, coupled with the small sample sizes (observations with monthly resolution go back to the middle of the 19th century) reduce its statistical power and it tends to fail for complex systems [120, 78]. Hence, it may mostly be applicable to bivariate analysis. Despite these drawbacks, we still employ the bivariate version of Granger Causality in this thesis to provide baseline for evaluating the more complex methods.

4.3 PCMCI+

The Peter and Clark Momentary Conditional Independence Plus (PCMCI+)[131] is an extension of the PCMCI [133] algorithm, which in itself is based on the Peter and Clark (PC)[141] algorithm. To provide a better explanation of PCMCI+ and why it is useful, it is a good idea to first explain PC, as it is the first step in the PCMCI and PCMCI+ algorithms. We are using the `tigramite` [132] Python package for running the PCMCI+ algorithm.

4.3.1 PC

The PC algorithm [141] is a causal discovery method based on conditional independence (CI) tests, meaning that we determine a variable X is **not** causal to Y given a conditioning set S, containing other variables in the system.

$$X \perp\!\!\!\perp Y|S \tag{4.4}$$

In probabilistic terms, it means that the **joint** probability of X and Y, given the conditioning set S is equal to the product of X given S and Y given S. The set S introduces knowledge into the system, accounting for the dependence of X and Y, and telling us that any association of the two variables is entirely explained by S.

$$P(X, Y|S) = P(X|S)P(Y|S) \tag{4.5}$$

The PC algorithm operates on 4 assumptions:

1. Causal Markov Condition: Given that we know the direct cause of a variable, other variables in the causal graph do not provide any further information.
2. Faithfulness: Statistical independence between two variables X and Y implies d-separation in the graph, i.e. either there is no connection between X and Y or their path is blocked by a third variable.
3. No hidden confounders: The dataset contains all variables relevant to relationships between variables.
4. Acyclic: The causal graph has no cycles.

Before going into the details of the PC algorithm, I will make a tangible example to explain the concept of conditional independence. Say it's a hot summer day at the beach, the sun is shining without a single cloud in sight. Perfect weather to have some ice cream. Unfortunately, the store that sells the ice cream is a 20 minute walk away, so by the time you have come back, your ice cream has started to melt and you have gotten a sunburn as well. Only considering the melting ice cream and the sunburn, these events are **correlated** and we actually observed a dependence between them - my ice cream melted and then I got a sunburn. In probabilistic terms it means that the joint probability does **not** factorise into the marginal ones:

$$P(\text{sunburn}, \text{melting ice cream}) \neq P(\text{sunburn})P(\text{melting ice cream}) \tag{4.6}$$

To the reader, it is obvious that these events **correlate** but are not **causal**. Neither does your sunburn cause the ice cream to melt, nor does the melting ice cream give you a sunburn. Instead there is a so called **common driver**, which causes both events, namely the sun. When considering the sun as a conditioning event, the melting ice cream and sunburn become **conditionally independent**. Knowing that the sun was shining, provides extra information about the sunburn and melting ice cream events, leading us to the conclusion that

$$P(\text{sunburn}, \text{melting ice cream}|\text{sun}) = P(\text{sunburn}|\text{sun})P(\text{melting ice cream}|\text{sun}) \tag{4.7}$$

In a graphical representation of this system, we can nicely highlight the difference between the conditional and unconditional probabilities. In figure 4.5, we have just observed our sunburn and the melting ice cream (the sun and arrows to the other nodes are greyed out to show that they are unobserved), leading to the conclusion that these events are dependent, highlighted by the dotted line between the sunburn and melting ice cream node. When observing the sun, as shown in figure 4.6, we understand that it is the real cause behind the two events and the sunburn and melting ice cream are independent, given the knowledge of the sun.

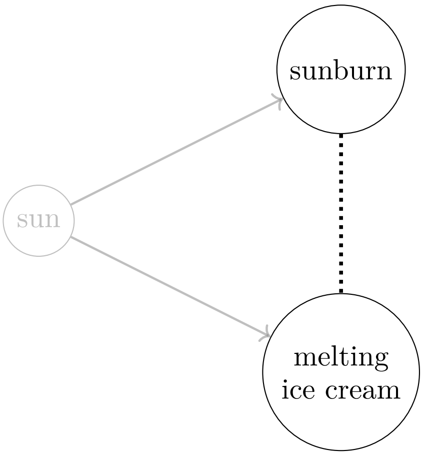


Figure 4.5 Observing just the sunburn and melting ice cream, we think they are dependent, as shown by the dotted line.

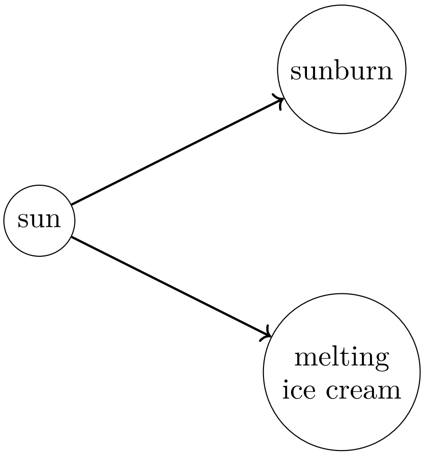


Figure 4.6 Observing the sun, sunburn and melting ice cream, we see that the sun is a common cause for both sunburn and melting ice cream.

In this example, it is very easy to tell which events are causal to others. When dealing with observational data, we instead need to run statistical tests to quantify the relationships. Using partial correlation for example, we can check whether removing the effect of a third variable Z, changes the correlation between X and Y. If the correlation vanishes, we can consider that their correlation is explained by Z.

The algorithm infers a causal structure of a set of variables by starting with a complete undirected graph, where all variables are connected, and progressively removing edges between variables based on Conditional Independence (CI) tests.

The algorithm follows a two-step process:

1. **Skeleton Identification:** Construct a fully connected undirected graph $G = (V, E)$ (Figure 4.7). The goal of this step is to go from a complete graph, to an undirected graph (Figure 4.8) representing the underlying causal model of the data, by iteratively removing edges. An edge $\{v_1, v_2\}$ can be removed in two situations: v_1 and v_2 are **independent**. If two variables are independent, they are by definition not causal to each other. If they are not independent, they might be conditionally independent, for example if there is a common driver. If we find them to be conditionally independent, we can also remove the edge.

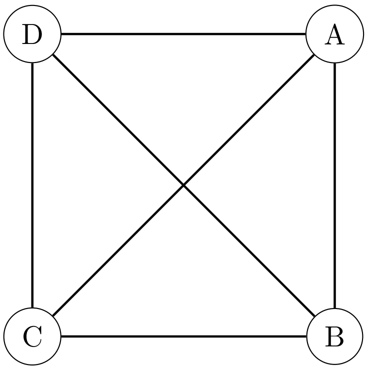


Figure 4.7 The initial complete graph where all variables are connected.

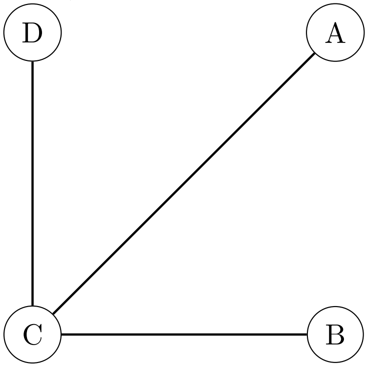


Figure 4.8 The undirected causal graph after removing edges.

2. **Orientation of Edges:** After obtaining the skeleton (figure 4.8), we have an undirected graph. Using the causal Markov property, faithfulness and acyclicity assumption, we want to orient the edges. Considering all *unshielded* triplets $A - C - B$, i.e. triplets of nodes where the edge $\{A, B\}$ was removed in the first phase and we have edges $\{A, C\}, \{B, C\}$, if the node C was **not** in the conditioning set when the edge $\{A, B\}$ was removed, then the edges from A and B can only point towards C. We can make this conclusion because if C was actually pointing to A and/or B, we would have measured a dependence between A and B.

Although the PC algorithm was not designed with timeseries data in mind, it can still work with them. Say variables X and Y are timeseries, then each node in the causal graph is an observation in the timeseries. The PC algorithm assumes causal sufficiency, meaning that there are no hidden confounders affecting the observed variables. While it is computationally efficient for small datasets, its performance deteriorates with increasing variable count due to the exponential growth in the number of CI tests.

4.3.2 PCMCI

In this section I will explain the workings of the PCMCI algorithm as defined by Runge et. al [133]. All the definitions and examples I give here are based on their work. When applying the PC algorithm to timeseries data, it suffers from large conditioning sets which actually only contain few relevant variables. This in turn unnecessarily increase the time complexity and reduces the algorithm's power to detect true causal links. PCMCI seeks to address this issue by reducing the conditioning set S to only include relevant variables. Thanks to the fact that the causal direction can only be forwards in time, we can already exclude all future values from the conditioning set. Additionally, we do not need to consider the entire past of a variable, conditioning on a set that contains at least the parents is enough due to the *Causal Markov Condition* [133].

The spatio-temporal structure of the data is best visualised with a *timeseries graph* (Figure 4.9). In the timeseries graph, nodes represent the variables at given points in time. Arrows between nodes indicate a causal link between nodes, where the direction of the arrow determines the flow of the relationship.

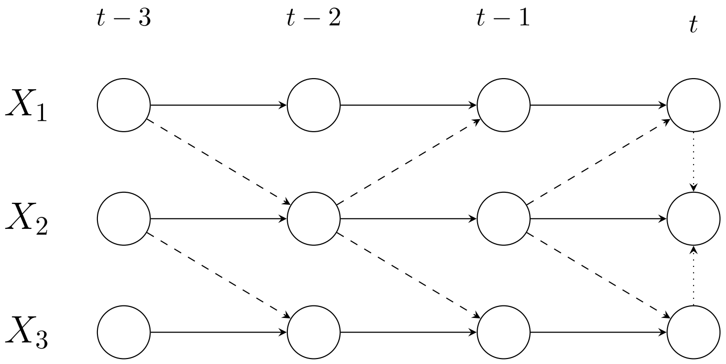


Figure 4.9 Example of a timeseries graph with variables X_1, X_2, X_3, show at 4 different time lags, $t-3, t-2, t-1, t$. Each row represents a single variable through time, columns show the 4 time lags. To better highlight the relationships in the graph, I have chosen different arrow line styles. This is purely a visualisation method, the condition for a link to exist is the same no matter which two nodes it connects. Filled arrows show **autoregressive** relationships, e.g. the edge $X_1^{t-3} \to X_1^{t-2}$. Note that links are not limited to a time lag $\tau = 1$, there could also be an edge $X_1^{t-3} \to X_1^t$. Dashed arrows represent relationships **between** variables in the graph, e.g. the edge $X_1^{t-3} \to X_1^t$, which goes from variable X_1 to variable X_2. Lastly, there are the **contemporaneous** relationships, shown by the dotted arrows, e.g. $X_1^t \to X_2^t$.

Remembering what we said about conditional independence in the PC algorithm section 4.3.1, an edge $X_i^{t-\tau} \to X_j^t$ exists if we **do not** measure any conditional independence between $X_i^{t-\tau}$ and X_j^t. If we were to measure any conditional independence, any statistical relationship between the two variables would be explained by the variables we condition on in set S and there would be no edge between the two. Not measuring any conditional independence means that there is **no other** variable in the system that could explain the statistical association we measured. Formally, we say that the edge $X_i^{t-\tau} \to X_j^t$ exists if:

$$X_i^{t-\tau} \not\perp\!\!\!\perp X_i^t | \mathbf{X}_-^t \setminus \left\{X_{t-\tau}^i\right\} \tag{4.8}$$

Explanation

- $X_i^{t-\tau}$: is the variable X_i at time lag $t-\tau$.
- X_j^t: is the variable X_j at time lag t.
- $\mathbf{X}_-^t \setminus \left\{X_{t-\tau}^i\right\}$: is the set of the past of **all** variables, up to a time lag $\tau_{\max}$, excluding the source variable $X_{t-\tau}^i$.

Meaning that the variable X_i at time lag $t-\tau$ and the variable X_j at time lag t are **not independent** when considering the set of the past of **all** variables, up to a time lag $\tau_{\max}$ excluding X_i. The problem here is that we include the past of **all** variables, which leads to large conditioning set S that reduces the effect size of the CI test [133]. Thanks to the *Causal Markov Property*, which says that given that we know the direct cause of a variable other variables in the graph do not provide further information, finding the parents is enough to establish a conditional independence. Hence, the goal is to find the parents of a node, such that we can condition on them. Luckily we know an algorithm that can do this for us, namely the PC algorithm. PCMCI employs a variation of the PC algorithm that is adapted to timeseries and seeks to condition on relevant parents only. Let $\mathcal{P}(X_j^t)$ be the set of preliminary parents of the variable X_j^t. At first, we initialise this to the past of **all** variables in the system $\mathcal{P}(X_j^t) = \mathbf{X}_-$. From this set, we measure associations between variables using a statistical test and remove all the **uncorrelated** variables, since *no correlation = no causation*, based on some significance threshold α. Then we rank the variables in $\mathcal{P}(X_j^t)$ according to their measured association to X_j^t, in descending order, such that variables with strongest association come first. Then we perform **conditional independence** tests, by conditioning on the variable with the strongest association. Any variables that are conditionally independent give this strongest driver are removed from $\mathcal{P}(X_j^t)$. Next, we condition on the two strongest drivers and remove any variables that become conditionally independent. We continue doing this, conditioning on the k strongest drivers at iteration k, until all strongest drivers have been exhausted. This leaves us with a conditioning set that includes the causal parents with high probability, but also some false positives since CI tests are not perfect, especially in finite samples, and may fail to detect true independencies due to noise or weak effect sizes. This is why PCMCI introduces a second phase - Momentary Conditional Independence (MCI) testing - to refine the parent set further by testing each candidate parent for conditional independence from X_j^t, this time conditioning on **all** other parents in the set (excluding $X_i^{t-\tau}$). Only variables that remain significantly associated under this stricter test are retained as final causal parents. We establish a link $X_i^{t-\tau} \rightarrow X_j^t t$ if and only if:

$$MCI : X_i^{t-\tau} \not\perp\!\!\!\perp X_j^t \quad | \quad \mathcal{P}(X_j^t) \setminus \left\{X_i^{t-\tau}\right\}, \mathcal{P}(X_i^{t-\tau}) \tag{4.9}$$

Explanation

To test if there is a link from variable X_j at time t to X_i at $t-\tau$, we measure if they are conditionally dependent. We condition on the set containing the parents of X_j at time t (with the variable $X_i^{t-\tau}$ removed), and the parents of X_i at time $t-\tau$. If we measure a conditional dependence, it means that the parents of both variables cannot sufficiently explain the measured relationship and so there must be a link between them.

4.3.3 PCMCI+

A drawback of PCMCI is that it does not discover contemporaneous relationships, where variable X_i influences variable X_j at time t. To achieve this, PCMCI+ splits the PC step into two phases: a *lagged* phase and a *contemporaneous* phase. The idea behind splitting the PC step is to reduce the dimensionality of the conditioning sets by first only finding lagged parents, and then using these together with contemporaneous links to refine the lagged parents and to remove any spurious contemporaneous edges. As a result of the smaller conditioning sets, we increase statistical power, reduce the number of tests, and improve recall.

In the lagged phase, PCMCI+ tests all variable pairs $(X^i_{t-\tau}, X^j_t)$ for $\tau > 0$, using conditional independence (CI) tests that condition only on the p strongest adjacencies of X^j_t, rather than all subsets of adjacents. The p strongest adjacencies are the variables that have the strongest statistical association, given by an **unconditional** independence test. The resulting set of lagged adjacents $\widehat{B}^-_t(X^j_t)$ contains both true lagged parents and spurious links due to paths through contemporaneous parents.

In the contemporaneous phase, PCMCI+ first initialises a graph where each node is connected to all other contemporaneous nodes S and to the lagged parents $\widehat{B}^-_t(X^j_t)$ found in the previous stage. Now, the method iteratively removes a link if the contemporaneous nodes and lagged parents cannot explain the statistical association that is measured between two variables $\left(X_i^{t-\tau}, X_t^j\right)$ using CI tests. The conditioning set for this CI test is therefore the parents $\widehat{B}^-_t(X^j_t)$ (excluding the past node $\widehat{B}^{t-\tau}_-(X_i^{t-\tau})$, $\widehat{B}^-_t(X^i_t)$, as well as the set of contemporaneous nodes S. We remove a link between $\left(X_i^{t-\tau}, X_t^j\right)$ if and only if:

$$MCI : X_i^{t-\tau} \perp\!\!\!\perp X_j^t \quad | \quad S, \widehat{B}^t_-(X^t_j) \setminus \left\{X_i^{t-\tau}\right\}, \widehat{B}^{t-\tau}_-(X_i^{t-\tau}) \tag{4.10}$$

Explanation

- S: the set of contemporaneous links.
- $\widehat{B}^t_-(X^t_j) \setminus \{X^{t-\tau}_i\}$: the lagged parents of X_j, without $X^{t-\tau}_i$.
- $\widehat{B}^{t-\tau}_-(X^{t-\tau}_i$: the lagged parents of $X^{t-\tau}_i$.

In PCMCI+, to test if we can remove the link from variable X_i at time $t-\tau$ to X_j at t, we measure if they are conditionally dependent. Like for PCMCI, we condition on the set containing the parents of X_j at time t (with the variable $X^{t-\tau}_i$ removed) $\widehat{B}^t_-(X^t_j)$, and the parents of X_i at time $t-\tau$, $\widehat{B}^{t-\tau}_-(X^{t-\tau}_i$. The difference to PCMCI, is that here we include also a set of **contemporaneous** links, S. If we do measure a conditional independence, it means that the link between the two variables can be explained by the parents of both variables found in the lagged phase and the contemporaneous links in the set S, and hence we can **remove** the link.

For our tests, we use the linear version of PCMCI+ that uses the partial correlation test. The partial correlation $ParCorr(X,Y|Z)$ is estimated with the correlation of the residuals of an ordinary least squares (OLS) fit of the two assumed models, see supplementary material in [133]:

$$X = \mathbf{Z}\beta_X + \epsilon \tag{4.11}$$

$$Y = \mathbf{Z}\beta_Y + \epsilon \tag{4.12}$$

Where the estimation of the correlation $\widehat{\rho}(\widehat{r_X}, \widehat{r_Y})$ uses:

$$\widehat{r_X} = X - \mathbf{Z}\widehat{\beta_X} \tag{4.13}$$

$$\widehat{r_Y} = Y - \mathbf{Z}\widehat{\beta_Y} \tag{4.14}$$

4.4 LKIF

Liang Kleeman Information Flow (LKIF) is a method based on the theory of information flow. In information flow, we study the transfer of information from a source variable x to a target variable y, where both variables are part of a dynamical system, often modelled as timeseries. Information flow found widespread application, in neuroscience [13, 29, 126, 166], analysing financial markets [30, 82, 94], social networks [24, 50, 172], and in atmosphere-ocean science [85, 32, 35, 118]. Specifically the Earth Science lends itself well since the atmospheric-ocean processes are part of a complex system which we always study through time.

Liang and Kleeman view causality as a real physical notion from which they derived their definition, first for 2D [88] and later for multivariate systems [86]. Specifically, LKIF operates within the context of a dynamical system governed by linear stochastic differential equations (SDEs). These equations describe how the system evolves over time, with the stochasticity introduced by a Wiener process, also known as Brownian motion. It is important to highlight that Liang assumed the system to be linear, which can be a good approximation in many situations but might not work when the system is too complex. Within this system, the goal of LKIF is to determine the **rate** at which information is transferred between different components, which we consider to be an indicator of causal relationships among them. In this context, information is quantified using the concept of entropy, which measures the amount of randomness or uncertainty present in the system. Hence, LKIF seeks to discover how the entropy of one component in the system changes due to the influence of another component and defines this as the rate of information transfer. Essentially, a high rate of information transfer indicates that the entropy of a component is significantly affected by another component. This influence is indicative of a causal relationship, where changes in one component lead to changes in the uncertainty or randomness of another. For a detailed derivation of LKIF, I refer the reader to the works [87] and [86], here I will merely show the final equations. The information flow from variable j to variable i is defined as follows:

$$T_{i\to j} = -\int\limits_{\mathbb{R}^d} \rho_{i|j}(x_j|x_i)\frac{\partial(F_i\rho_{\not j})}{\partial x_i}d\mathbf{x} + \frac{1}{2}\int\limits_{\mathbb{R}^d} \rho_{i|j}(x_j|x_i)\frac{\partial^2(g_{ii}\rho_{\not j})}{\partial x_i^2}d\mathbf{x} \tag{4.15}$$

Explanation

- $\rho_{i|j}(x_j|x_i)$: is the conditional probability density function of x_j conditioned on x_i.
- $\rho_{\not j} = \int\limits_{\mathbb{R}} \rho(\mathbf{x})dx_j$: is the marginalisation of the joint probability density function by integrating out the variable x_j.
- $g_{ii} = \sum\limits_{k=1}^{n} b_{ik}b_{ik}$: strength of the noise affecting variable x_i, where $b_{i,k}$ is the amplitude of the noise from the i-th Wiener process acting on the k-th variable.

For a detailed explanation and derivation I again refer the reader to [86]. This problem with this definition is that it requires the probability density function ρ, which we do not have in our scenario. Hence we need to estimate the information flow to have an easy to use formula that we can actually work with. Liang [86] proposes the following maximum likelihood estimator (MLE), derived from Equation 4.15:

$$\hat{T}_{i\to j} = \frac{1}{det\mathbf{C}}\sum_{k=1}^{d}\Delta_{jk}C_{k,di}\frac{C_{ij}}{C_{ii}} \tag{4.16}$$

Explanation

- $det\mathbf{C}$: Determinant of the covariance matrix that measures data correlation and concentration.
- d: The number of variables in the system.
- Δ_{jk}: Cofactors of the covariance matrix that measure a variable's influence on correlation and variance.
- $C_{k,di}$: Sample covariance between x_k and Euler approximation of $\frac{dx_i}{dt}$ tells us how x correlates with the rate of change of x_i.
- $\frac{C_{ij}}{C_ii}$: The sample covariance C_{ij} between x_i and x_j, normalised by the sample variance C_{ii} of x_i.

which is normalised with the term

$$Z_i = \sum_{k=1}^{d} |T_{k\to i}| + \left|\frac{dH_i^{noise}}{dt}\right| \tag{4.17}$$

where

Explanation

- $|T_{k\to i}|$: the information flow from all other variables to variable i.
- $\left|\frac{dH_i^{noise}}{dt}\right|$: the rate of change of the noise for variable i.

the final normalisation being as follows:

$$\tau_{i\to j} = \frac{\hat{T}_{i\to j}}{Z_i} \tag{4.18}$$

The Maximum Likelihood Estimation (MLE) in equation 4.16 utilizes various sample covariance measures along with the influence of variables on the overall variance and system dynamics to estimate the rate of information flow. The normalizer in equation 4.17 incorporates the information flow from all other variables as well as a noise component. This noise component accounts for the stochasticity within the system, ensuring that the information flow is measured in a way that reflects the inherent randomness and variability. By default LKIF measures the rate of information flow for contemporaneous relationships. To discover time lagged relationships, we have to add variables that are the shifted versions of the original timeseries. For instance, given two timeseries X, and Y, shifting X backwards by i lags means to align the two timeseries in way where X_{t-i} is aligned with Y_t. Essentially, this boils down to **removing** the **last** i entries in X and the **first** i entries in Y.

$$\begin{aligned} \text{lag}(X, 2) = X_{1:t-2} &= [x_1, x_2, \dots x_{t-2}] && (4.19)\\ Y_{3:t} &= [y_3, y_4, \dots y_t] && (4.20)\end{aligned}$$

Explanation

Given two timeseries with t elements, the $\text{lag}(X, i)$ function indexes the timeseries X from the first element 1 to element $t - i$. To match the length of this lagged timeseries, we also need to remove entries from Y, so we index from i to t.

Since we are working with an estimator, the point estimate of the information flow $\tau_{i \to j}$ alone is insufficient; it is necessary to assess its statistical significance. [34, 33, 161] employ bootstrap methods to approximate the sampling distribution of $\tau_{i \to j}$. In the bootstrap procedure, we generate b resampled datasets by drawing, with replacement, full-length time series from the original dataset for each variable. Each bootstrap sample retains the original temporal structure and is represented as $\mathbb{R}^{t \times n}$, where t is the number of time steps and n is the number of variables. For each of the b samples, the information transfer is recalculated, resulting in an empirical information flow distribution. The standard deviation of these estimates serves as an empirical measure of the standard error, which is used to construct the confidence interval under the assumption of a normal distribution at significance level α.

$$CI(\tau_{i \to j}) = \hat{\tau}_{i \to j} \pm z_\alpha \times SE(\hat{\tau}_{i \to j}) \tag{4.21}$$

Specifically, if the confidence interval includes zero, it suggests that despite the observed estimate, the true value of $\tau_{i \to j}$ may plausibly be zero, and thus the information flow cannot be considered statistically significant. We are using the LKIF Python implementation by Docquier [36].

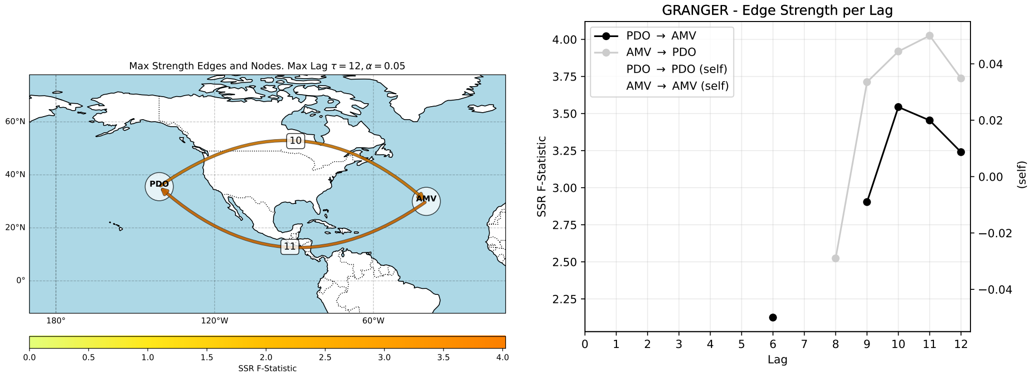
Chapter 5

Results

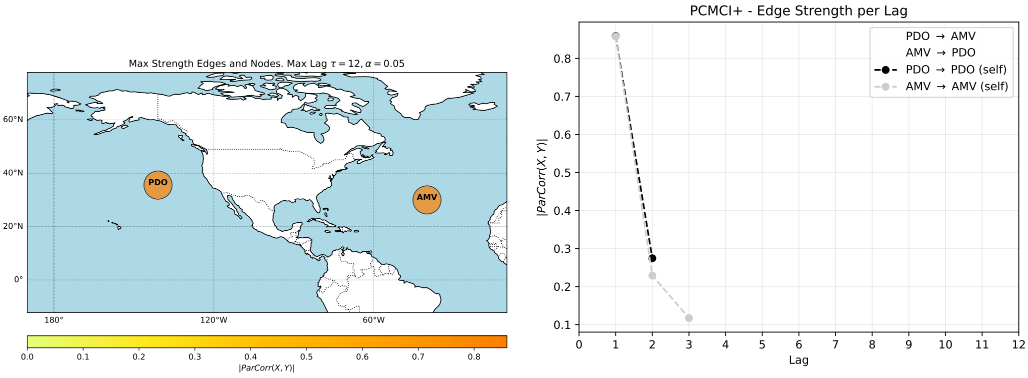
In the following section I will explain the results from running the three causal methods: Granger Causality, PCMCI+, LKIF, on the **smoothed** versions of the AMV and PDO index, as well as the **smoothed** varimax rotated principal components in the Pacific. We computed the causal relationships with lags The results are obtained with maximum lag, $\tau = 12$ at $\alpha = 0.05$ significance. The code and scripts used to produce the figures and results in this thesis are available at: `https://gitlab.com/sapienza-data-science-msc/thesis/data-exploration`.

5.1 AMV to PDO

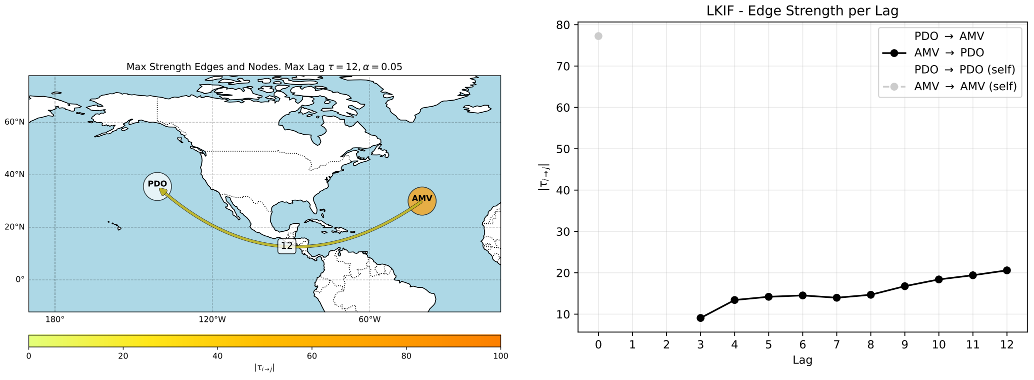
Granger causality finds links from AMV to PDO and vice versa, meaning that according to this method they two indices influence each other. The influence operates between lags 8 and 12, with a weak influence from the PDO to the AMV at lag 6. The strongest links are found at lag 11 and 10 as seen in Figure 5.1b, when the AMV is 11 months behind the PDO and the PDO is 10 months behind the AMV. Importantly, the connection from the AMV to the PDO peaks higher than the connection from PDO to AMV, showing that the influence from AMV to PDO is stronger than the other way around. PCMCI+ does not discover any links between PDO and AMV, which could be due to the fact that the method struggles in systems with two variables, something in agreement with Docquier et al. [34]. Hence, the map in Figure 5.1c does not contain any arrows and 5.1d is empty. Figure 5.1e shows that the information flow AMV to PDO increases with the lag (except for a small trough at lag 7) according to LKIF. There are however no links from PDO to AMV. The strongest link found by LKIF from AMV to PDO is at lag 12. Additionally, there is also no self information transfer for the PDO.



(a) Granger Causality **(b)** Granger Causality



(c) PCMCI+ **(d)** PCMCI+



(e) LKIF **(f)** LKIF

Figure 5.1 The graphs discovered by Granger Causality, PCMCI+, and LKIF when applied to the system of AMV and PDO and the edge strength per lag for the edges in the graph.

5.2 AMV to PDO and First 3 Varimax Principal Components

Here we see the results for a system consisting of the AMV, PDO, and the first three varimax rotated principal components. The graph found by Granger Causality is densely connected, with the strongest outgoing edges from the AMV being to the second principal component (SEPM), the third principal component (SPDOL) and first principal component (ENSO) (see Figure 5.3a). The link to the PDO is weaker than other outgoing edges from the AMV, it gains some strength from the lag 8 to lag 12. In fact, the edge from AMV to PDO in Figure 5.2 is already present in Figure 5.1a, where the system just contained the AMV and PDO timeseries. This is to be expected, since this bivariate version of Granger Causality always considers timeseries pairs in isolation and any other variables in the system do not affect whether the method finds a link or not. The Pacific region itself, i.e. nodes PDO, ENSO, SEPM, SPDOL, is also well interconnected. Considering the edge strengths in the entire graph across all lags, the outgoing links from AMV to SEPM and SPDOL are the absolute strongest. On the incoming edges to PDO, the strongest is coming from ENSO, followed by SPDOL, SEPM, and finally the AMV (Figure 5.3d).

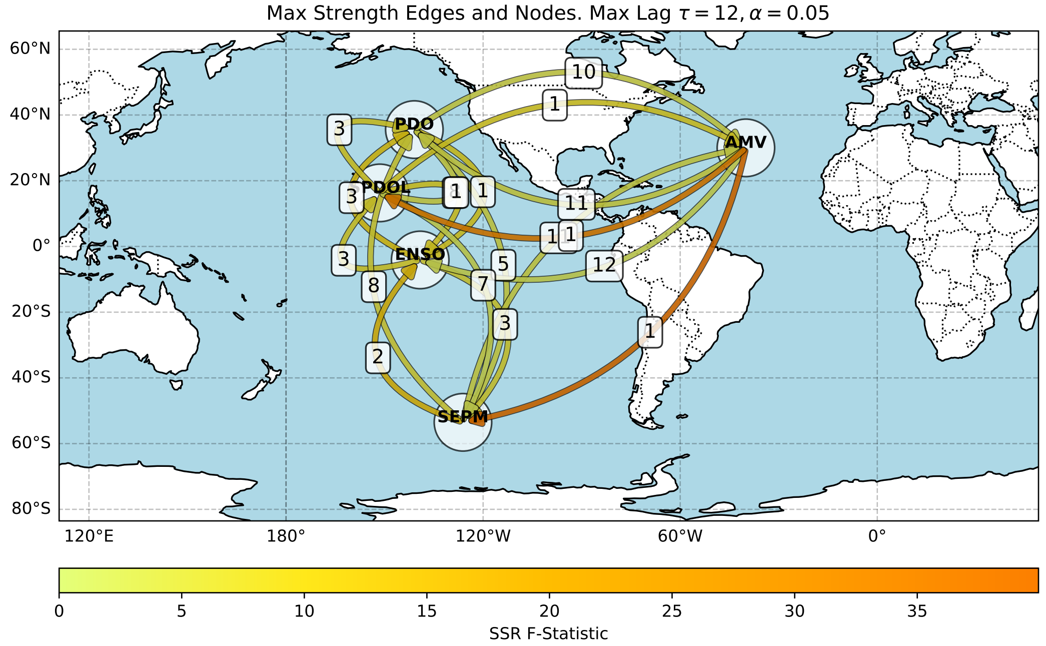
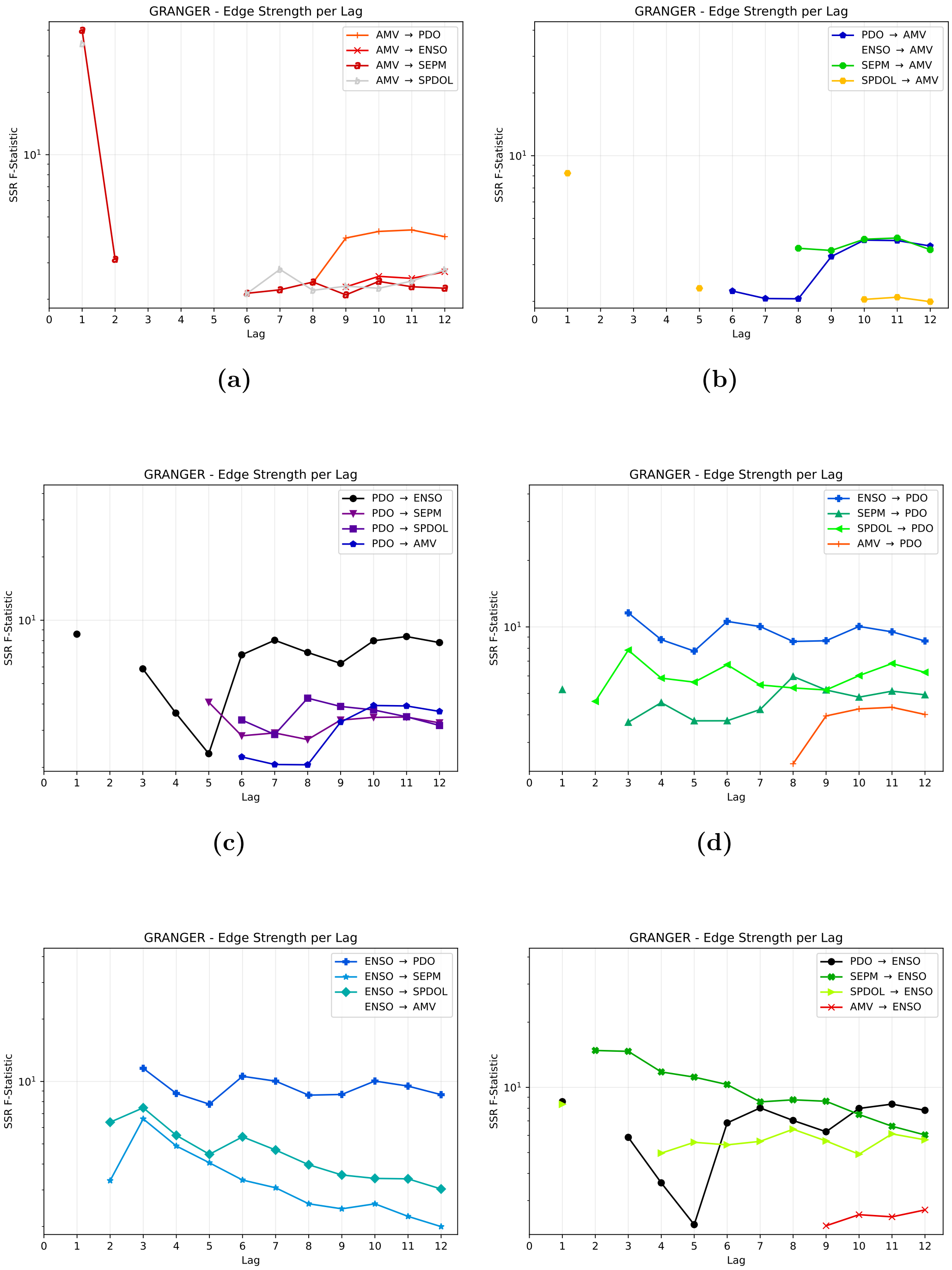


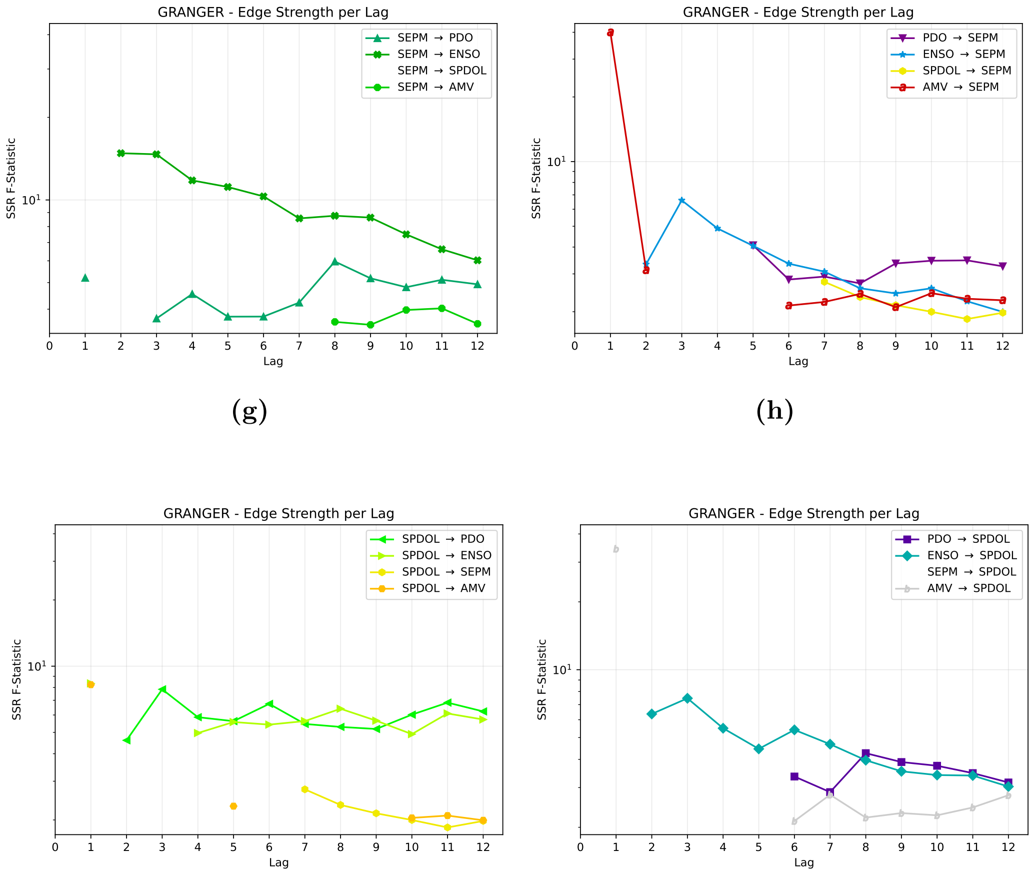
Figure 5.2 This plot shows the strongest links that Granger Causality found for the system containing the AMV and PDO indices and the first three varimax-rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the lag at which the link is at its strongest. Edge colour shows the F-statistic for two variables considered by the method.


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
AMV → PDO
AMV → ENSO
AMV → SEPM
AMV → SPDOL
PDO → AMV
ENSO → AMV
SEPM → AMV
SPDOL → AMV
PDO → ENSO
PDO → SEPM
PDO → SPDOL
PDO → AMV
ENSO → PDO
SEPM → PDO
SPDOL → PDO
AMV → PDO
ENSO → PDO
ENSO → SEPM
ENSO → SPDOL
ENSO → AMV
PDO → ENSO
SEPM → ENSO
SPDOL → ENSO
AMV → ENSO

(a) (b)

(c) (d)

(e) (f)



(g) **(h)**

(i) **(j)**

Figure 5.3 The edge strengths per lag for Granger Causality. Each panel shows the strength at each lag for outgoing or incoming links for a single node in the graph.

PCMCI+ finds contemporaneous links from AMV to SPDOL and vice versa (Figure 5.4). It also finds contemporaneous links from ENSO to PDO and vice versa. The strongest links are between ENSO and PDO, the connections from AMV and SPDOL are weaker. The auto *ParrCorr*, i.e. the partial correlation of the timeseries with its own past, is also the strongest at lag 1 for all nodes, as seen by the dashed line in figures 5.5. There are no direct links from AMV to PDO, which is the same as when we considered a system consisting of only the AMV and PDO in Figure 5.1c.

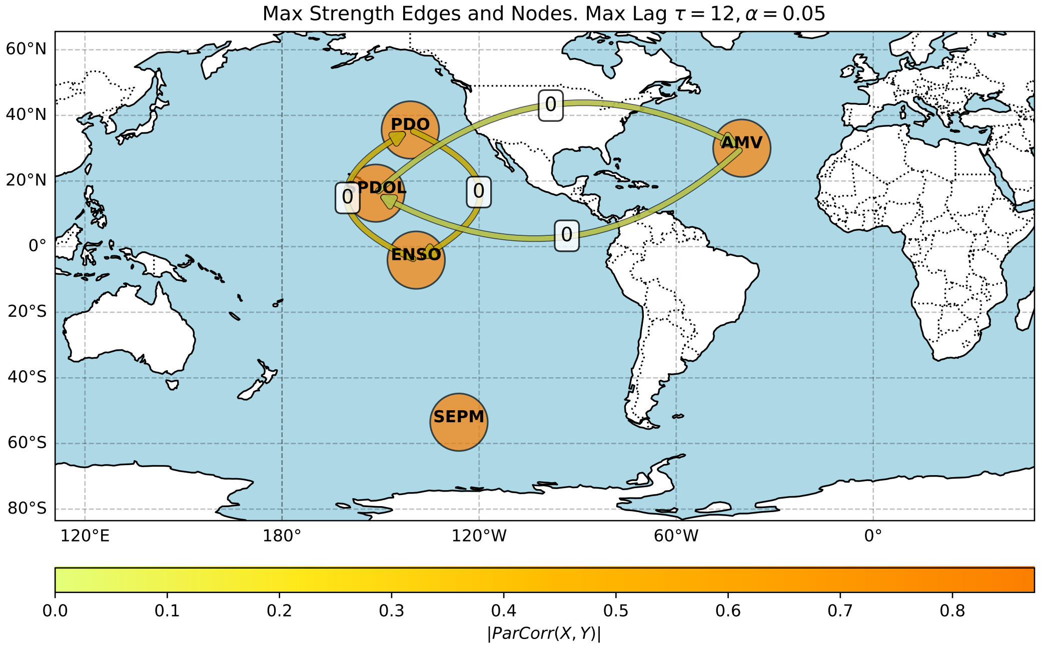
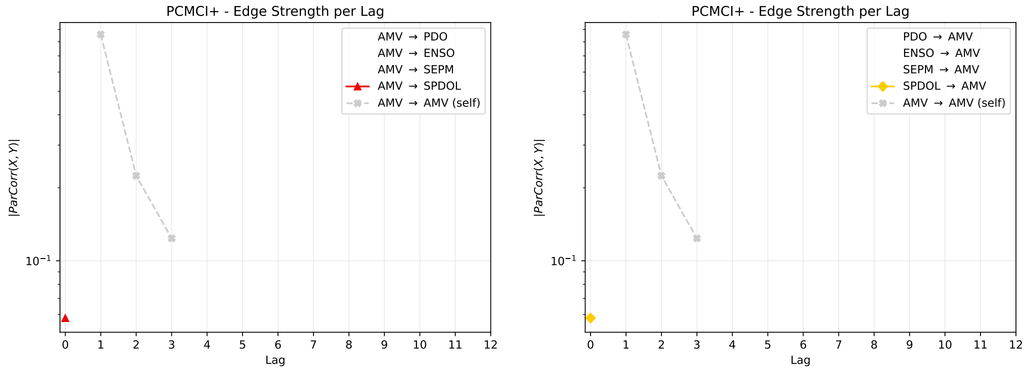
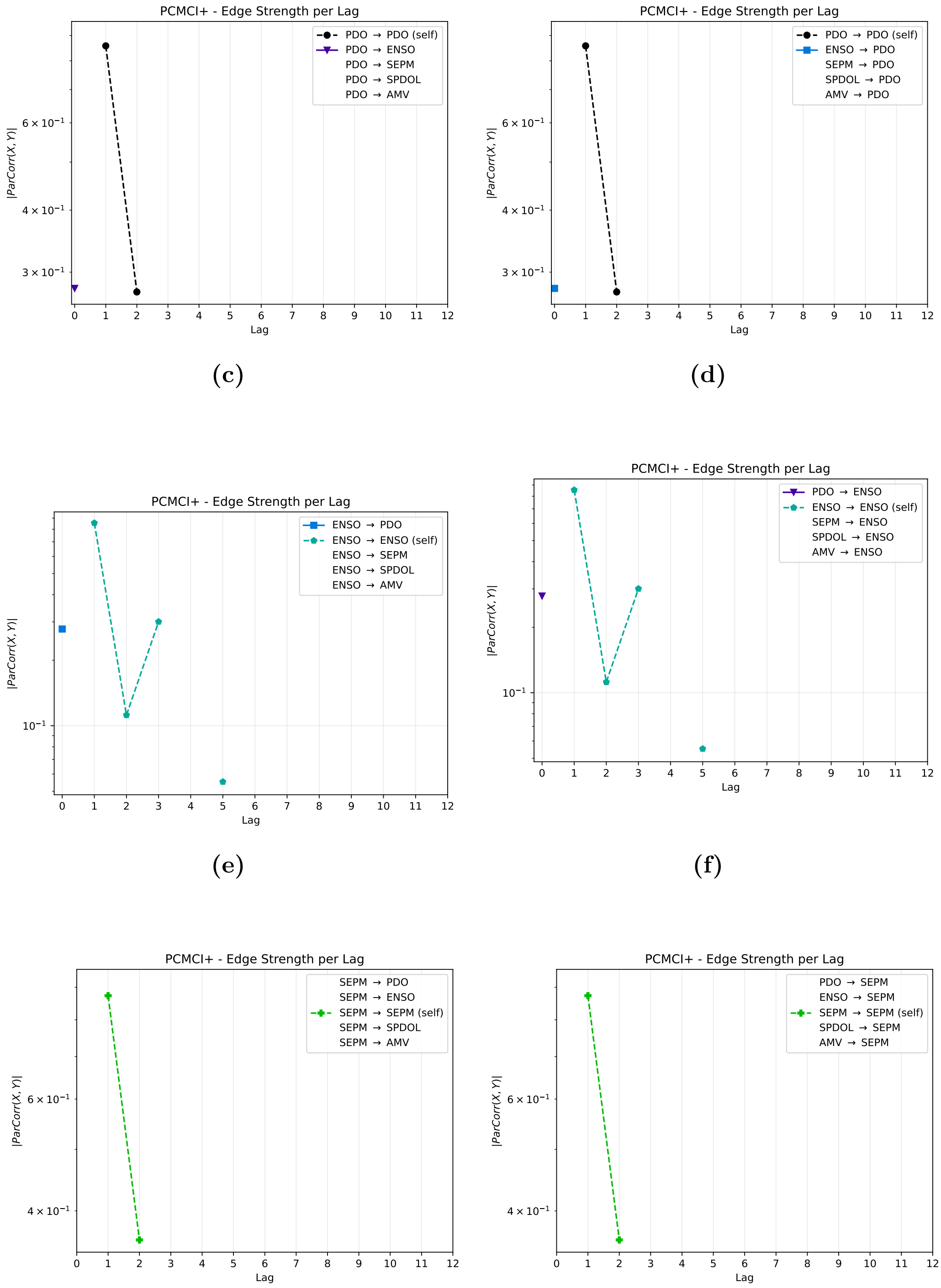


Figure 5.4 The strongest links found by PCMCI+ when considering a system with AMV, PDO, and the first three varimax rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the lag at which the link is at its strongest. Edge colour indicates the absolute value of the partial correlation between two variables. Node colour indicates the absolute value of the partial correlation of a variable and its own past.



(a) **(b)**

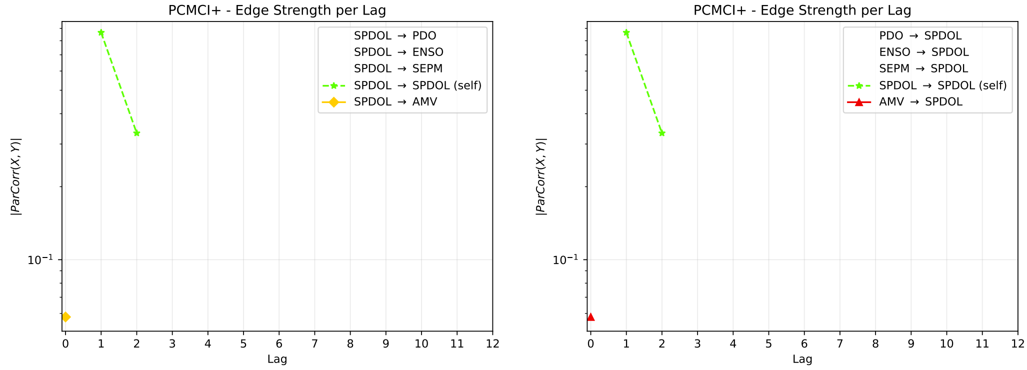


(c) (d)

(e) (f)

(g) (h)

LKIF on the other hand finds strong instantaneous links from the AMV to ENSO, SEPM, SPDOL, and the PDO, where the link to the PDO is the weakest among the four, see Figure 5.5a. The Pacific is, again, fairly interconnected, with a strong link from ENSO to PDO and weaker links from ENSO to SEPM and SPDOL, Figure 5.7e. An important aspect to remember when looking at the plots is that it is



(i) (j)

Figure 5.5 The edge strengths per lag for PCMCI+.

only the AMV timeseries which is lagged. The other timeseries remain unchanged, hence the influence among any timeseries that is not the AMV, is only calculated for the contemporaneous lag. For example, the influence from ENSO to PDO in Figure 5.7d is always calculated at lag 0, the y-axis lag refers to the AMV.

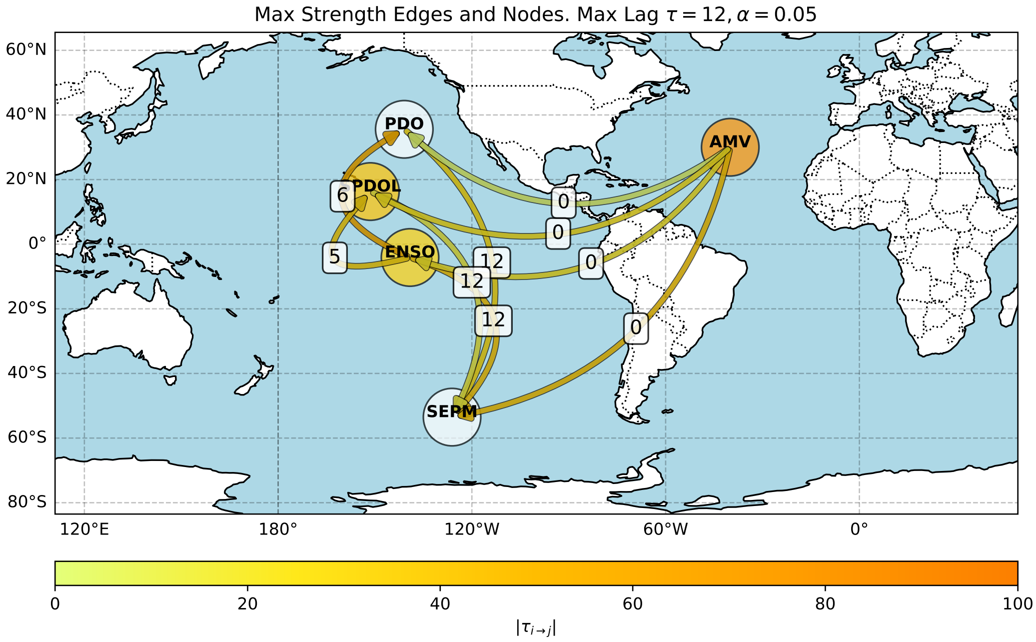
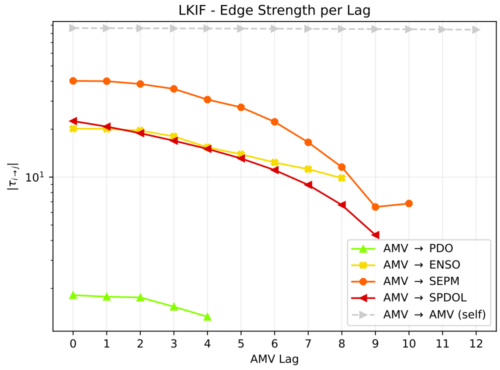
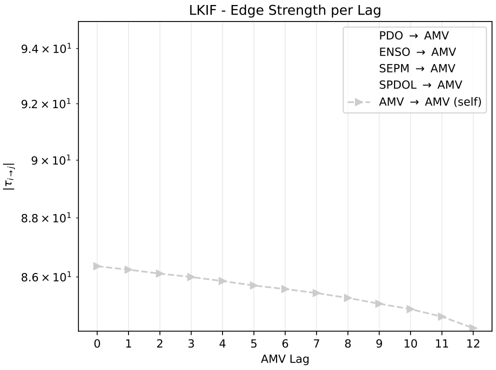


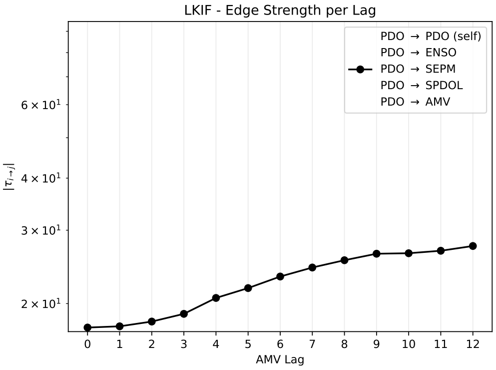
Figure 5.6 The strongest links found by LKIF when considering a system with AMV, PDO, and the first three varimax rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the AMV lag at which the link is at its strongest. Edge colour indicates the absolute value of the normalised information flow. Node colour indicates the absolute value of the normalised information flow of a variable onto itself.


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
AMV → PDO
AMV → ENSO
AMV → SEPM
AMV → SPDOL
AMV → AMV (self)

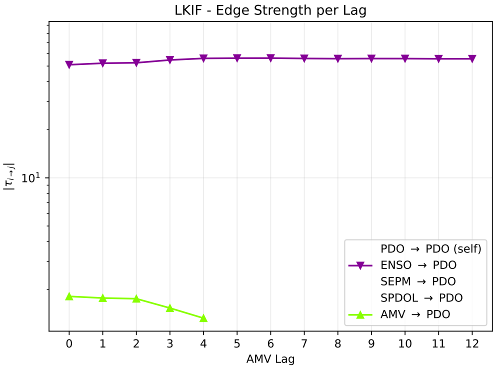
(a)


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
9.4×10^1
9.2×10^1
9×10^1
8.8×10^1
8.6×10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
PDO → AMV
ENSO → AMV
SEPM → AMV
SPDOL → AMV
AMV → AMV (self)

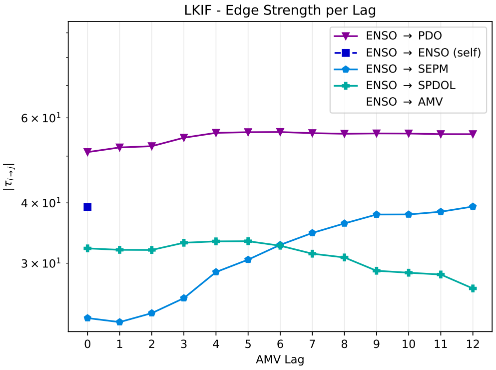
(b)


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
6×10^1
4×10^1
3×10^1
2×10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
PDO → PDO (self)
PDO → ENSO
PDO → SEPM
PDO → SPDOL
PDO → AMV

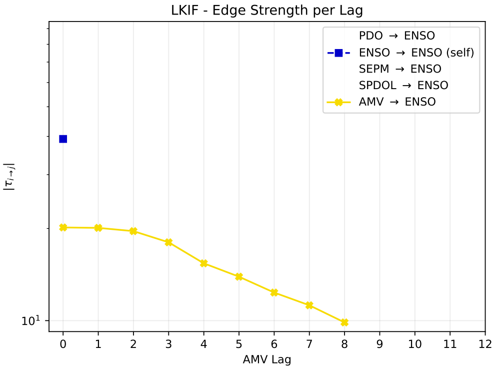
(c)


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
PDO → PDO (self)
ENSO → PDO
SEPM → PDO
SPDOL → PDO
AMV → PDO

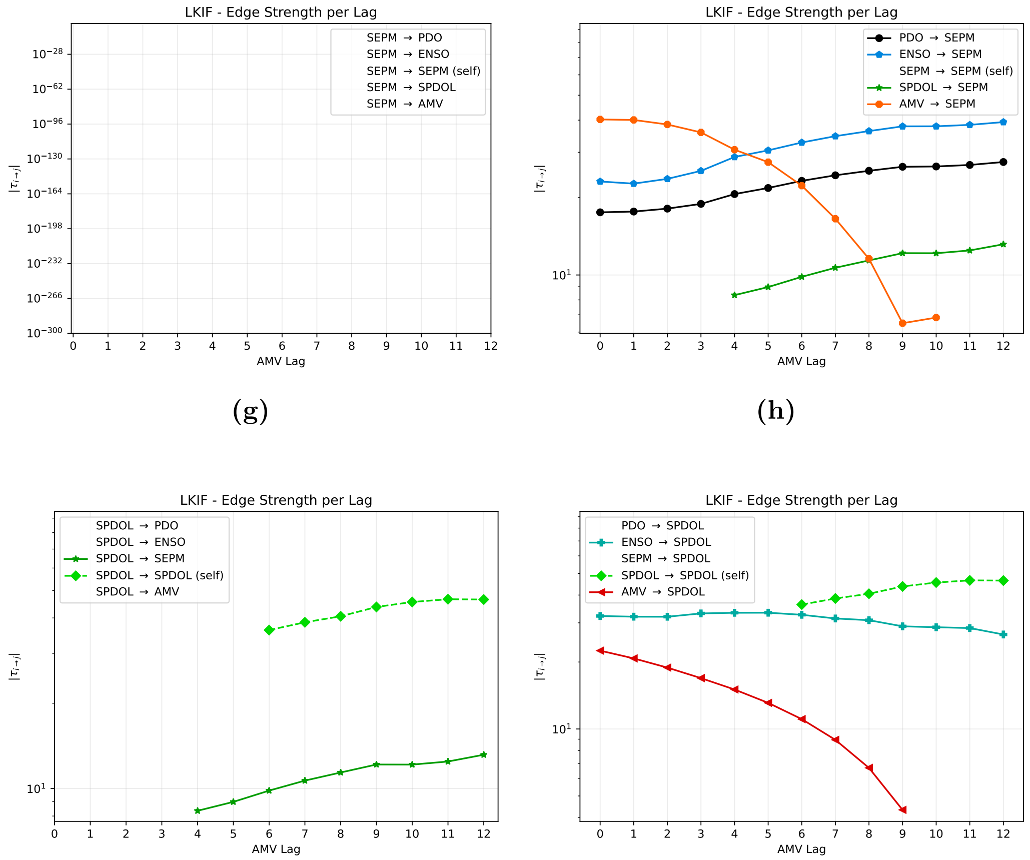
(d)


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
6×10^1
4×10^1
3×10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
ENSO → PDO
ENSO → ENSO (self)
ENSO → SEPM
ENSO → SPDOL
ENSO → AMV

(e)


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
10^1
AMV Lag
0
1
2
3
4
5
6
7
8
9
10
11
12
PDO → ENSO
ENSO → ENSO (self)
SEPM → ENSO
SPDOL → ENSO
AMV → ENSO

(f)



(g) **(h)**

(i) **(j)**

Figure 5.7 The edge strengths per AMV lag for LKIF. Each panel shows the strength at each lag for outgoing or incoming links for a single node in the graph.

5.3 AMV to PDO and First 5 Varimax Principal Components

Here we see the results for a system consisting of the AMV, PDO, and the first five varimax principal components. Here it is important to note that the due to the varimax rotation, the first three principal components in a system with $n = 3$ variables are **NOT** the same principal components as in a system with $n = 5$ components, since the rotation always rotates the entire system to optimise component sparsity. As with three components, Granger causality discovers a dense graph to the point where it is impossible to analyse visually. The strongest outgoing links from AMV are to the principal components SPD and SEPM, whereas the link to the PDO is much weaker, see Figure 5.9a. Then, looking at the incoming links to PDO, we observe that ENSO is the strongest 5.9d. The node that has the strongest incoming links is SPD 5.9h, with strong links coming from AMV, SPDOL, and PDO.

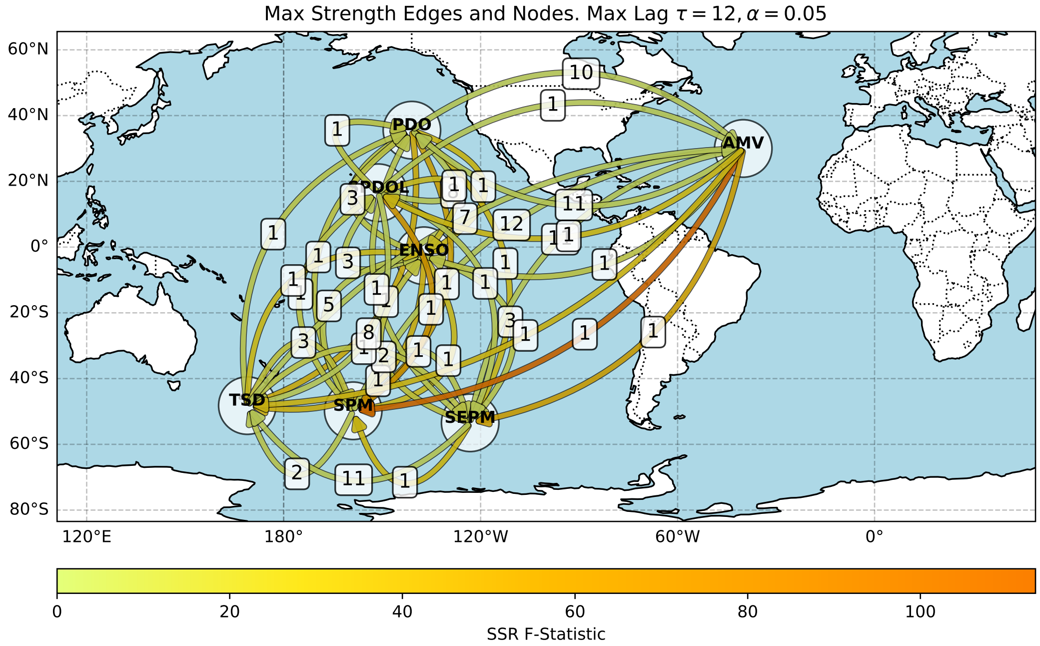
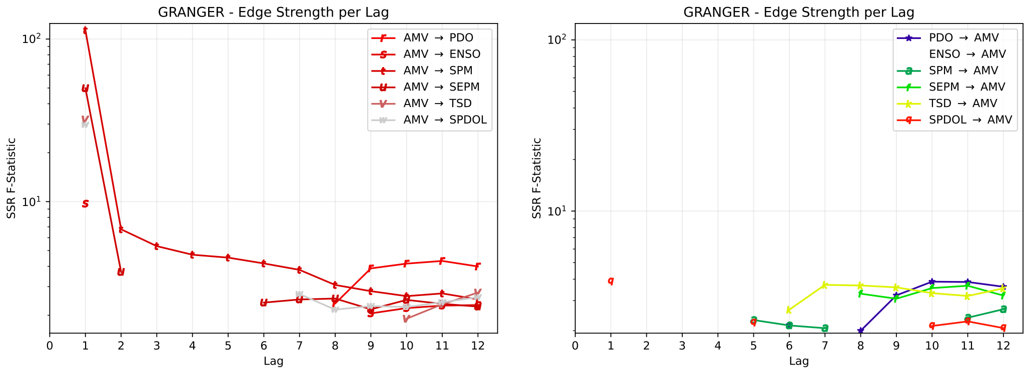
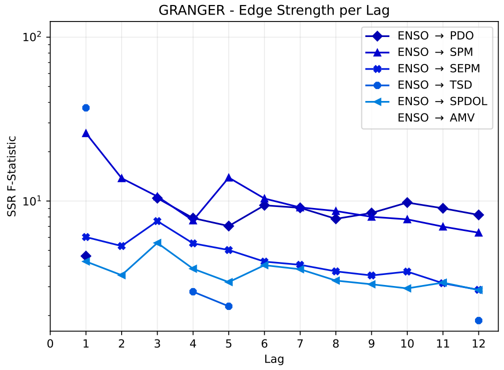


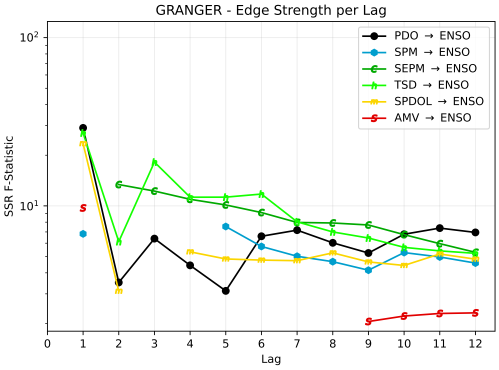
Figure 5.8 This plot shows the strongest links that Granger Causality found for the system containing the AMV and PDO indices and the first five varimax-rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the lag at which the link is at its strongest. Edge colour shows the F-statistic for two variables considered by the method.



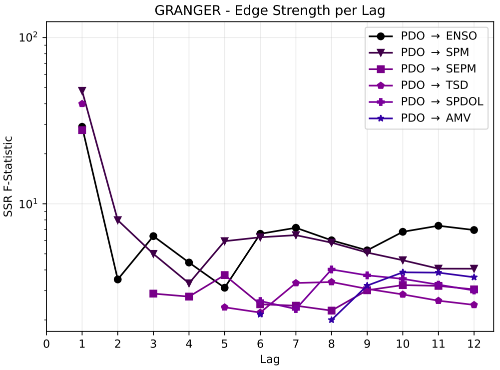
(a) (b)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
ENSO → PDO
ENSO → SPM
ENSO → SEPM
ENSO → TSD
ENSO → SPDOL
ENSO → AMV

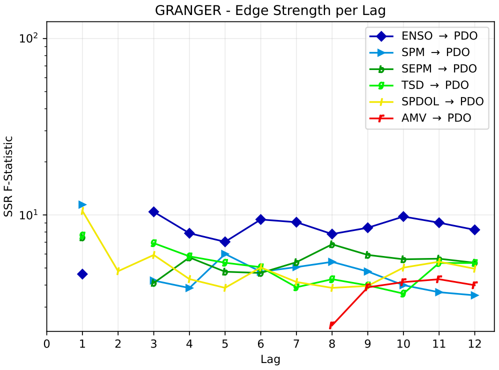
(e)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
PDO → ENSO
SPM → ENSO
SEPM → ENSO
TSD → ENSO
SPDOL → ENSO
AMV → ENSO

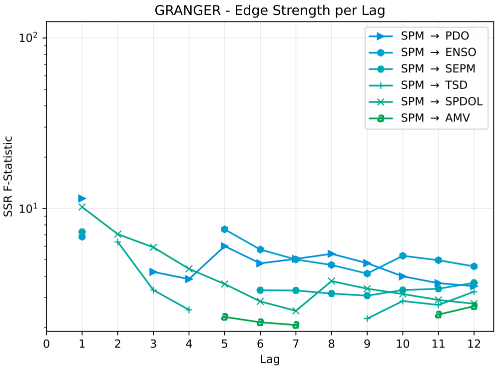
(f)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
PDO → ENSO
PDO → SPM
PDO → SEPM
PDO → TSD
PDO → SPDOL
PDO → AMV

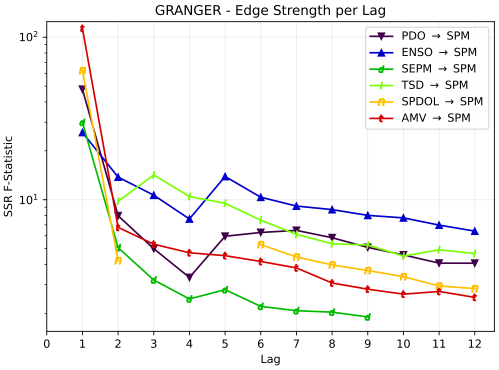
(c)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
ENSO → PDO
SPM → PDO
SEPM → PDO
TSD → PDO
SPDOL → PDO
AMV → PDO

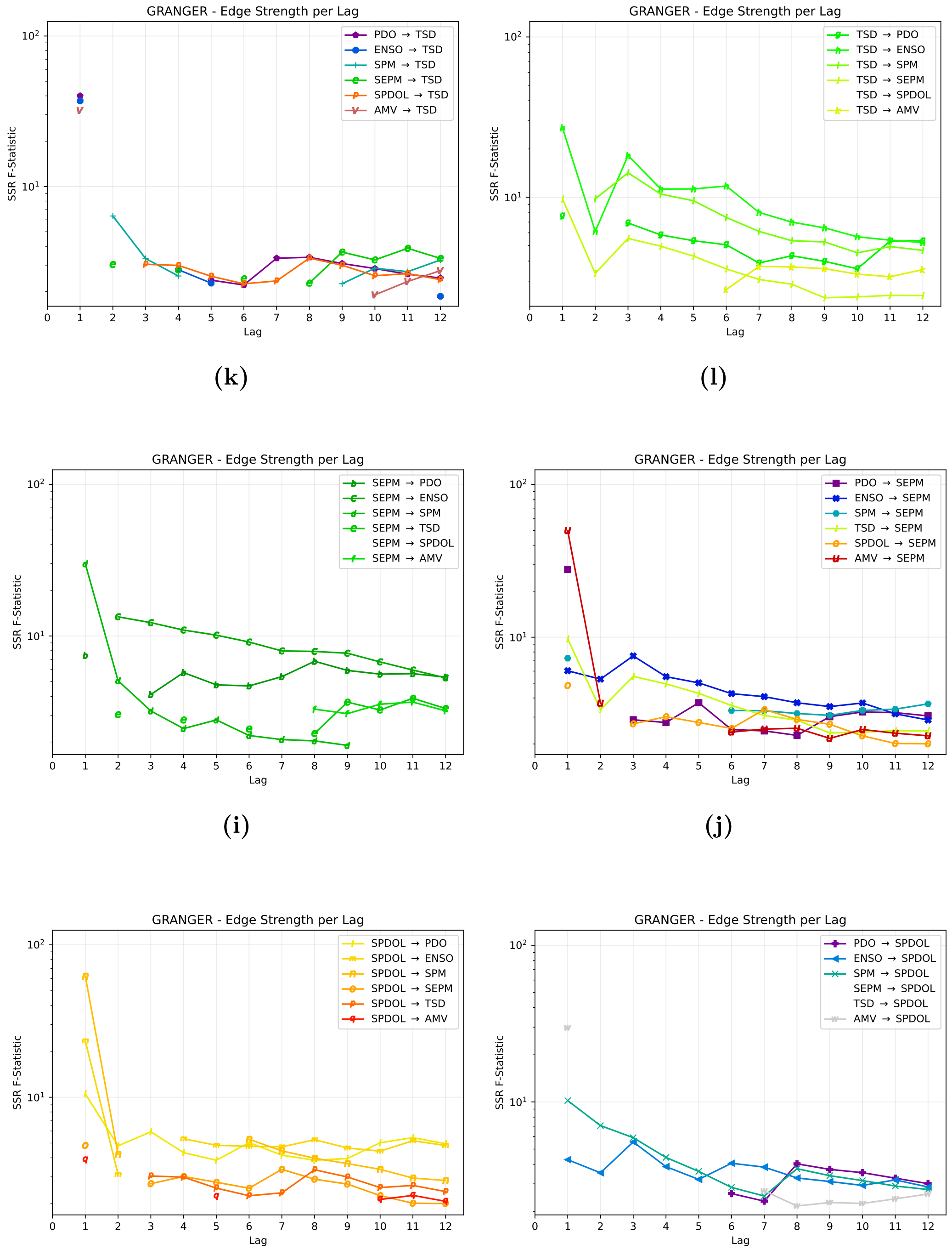
(d)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
SPM → PDO
SPM → ENSO
SPM → SEPM
SPM → TSD
SPM → SPDOL
SPM → AMV

(g)


GRANGER - Edge Strength per Lag
SSR F-Statistic
10^2
10^1
Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
PDO → SPM
ENSO → SPM
SEPM → SPM
TSD → SPM
SPDOL → SPM
AMV → SPM

(h)



(k) (l)

(i) (j)

(m) (n)

Figure 5.9 The edge strengths per lag for Granger Causality. Each panel shows the strength at each lag for outgoing or incoming links for a single node in the graph.

Interestingly, PCMCI+ finds only a single link in the system, namely from SPDOL to SPM. The plots below therefore only show the self influence of each node as well as the single link, seen in 5.11g and 5.11e. There are no links from AMV to any of the Pacific nodes, like we see for 3 varimax rotated components (5.4), which can be explained by the fact that the additional components explain enough of the statistical associations to render the links found previously, e.g. from AMV to ENSO, conditionally independent. We also see that the value of the self link (node colours in 5.11)is the same at lag 1 for all nodes (AMV 5.11a, PDO 5.11b, ENSO 5.11c, SPM 5.11e, SEPM 5.11d, TSD 5.11f, SPDOL 5.11g), which could be caused by the lack of parents that each node has. For PCMCI+, the self-link at lag 1 reflects the partial correlation between a variable and its own past - that is, $ParCorr(X_i^t, X_i^{t-1})$. In the absence of other predictors, this partial correlation essentially captures the autocorrelation structure of the time series itself. Smoothing - such as a running mean - artificially inflates autocorrelation by construction. Since each smoothed value is an average over past observations, adjacent values become more similar, increasing their correlation. This smoothing-induced autocorrelation might lead to the identical self-link strengths across nodes at lag 1.

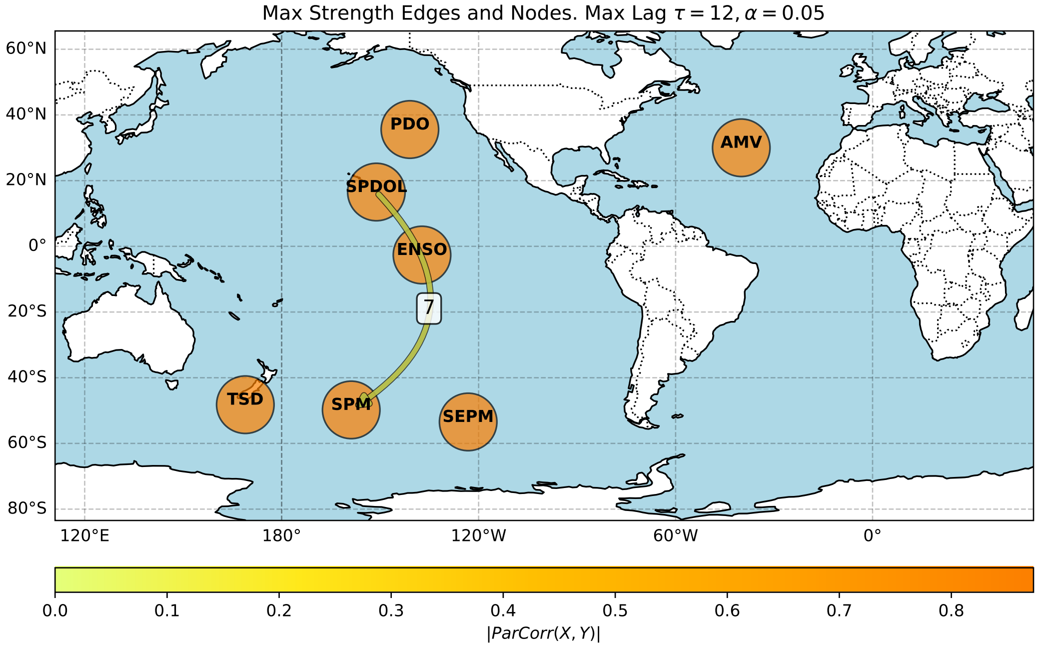
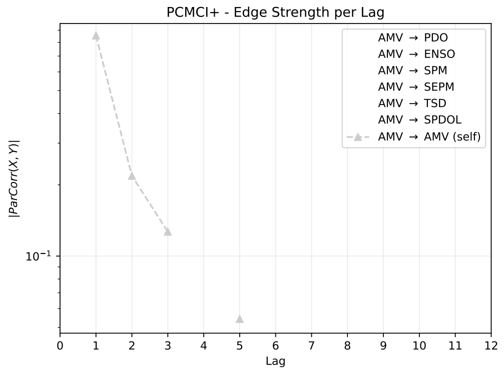
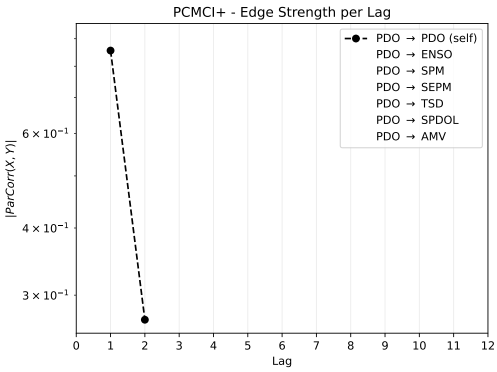


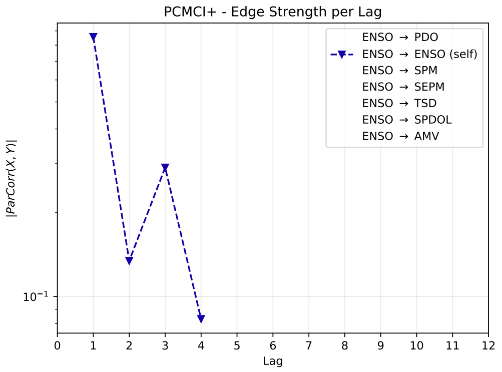
Figure 5.10 The strongest links found by PCMCI+ when considering a system with AMV, PDO, and the first five varimax rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the lag at which the link is at its strongest. Edge colour indicates the absolute value of the partial correlation between two variables. Node colour indicates the absolute value of the partial correlation of a variable and its own past.


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
AMV → PDO
AMV → ENSO
AMV → SPM
AMV → SEPM
AMV → TSD
AMV → SPDOL
AMV → AMV (self)

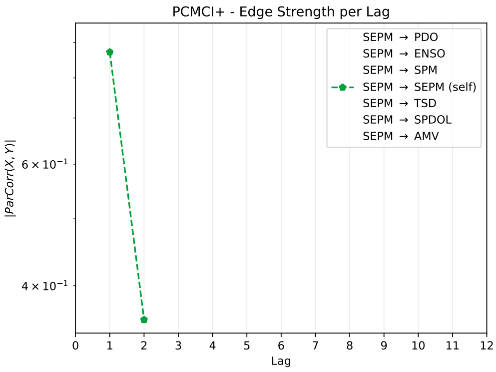
(a)


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
6×10^{-1}
4×10^{-1}
3×10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
PDO → PDO (self)
PDO → ENSO
PDO → SPM
PDO → SEPM
PDO → TSD
PDO → SPDOL
PDO → AMV

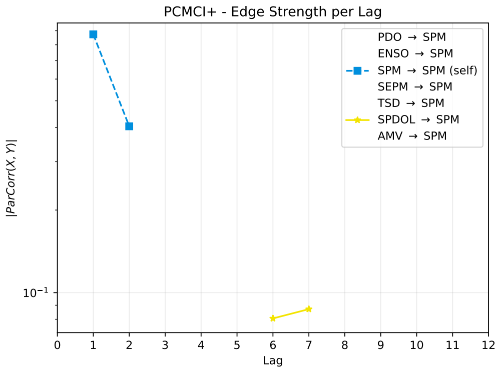
(b)


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
ENSO → PDO
ENSO → ENSO (self)
ENSO → SPM
ENSO → SEPM
ENSO → TSD
ENSO → SPDOL
ENSO → AMV

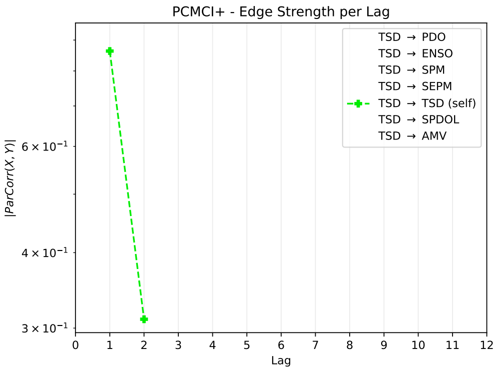
(c)


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
6×10^{-1}
4×10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
SEPM → PDO
SEPM → ENSO
SEPM → SPM
SEPM → SEPM (self)
SEPM → TSD
SEPM → SPDOL
SEPM → AMV

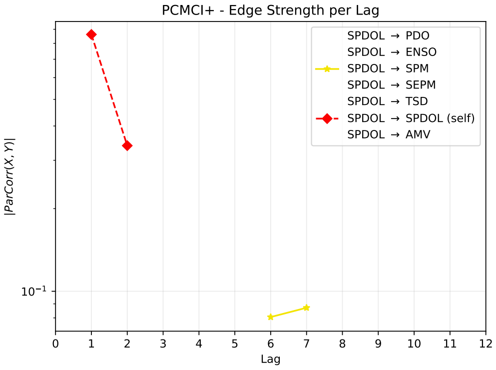
(d)


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
PDO → SPM
ENSO → SPM
SPM → SPM (self)
SEPM → SPM
TSD → SPM
SPDOL → SPM
AMV → SPM

(e)


PCMCI+ - Edge Strength per Lag
|ParCorr(X, Y)|
6×10^{-1}
4×10^{-1}
3×10^{-1}
0
1
2
3
4
5
6
7
8
9
10
11
12
Lag
TSD → PDO
TSD → ENSO
TSD → SPM
TSD → SEPM
TSD → TSD (self)
TSD → SPDOL
TSD → AMV

(f)



(g)

Figure 5.11 The edge strengths per lag for PCMCI+. Each panel shows the strength at each lag for outgoing or incoming links for a single node in the graph.

Similar what we already saw for three varimax components, LKIF only finds links from AMV to the Pacific (5.12), but not the other way around. The link from the AMV to the PDO is once again, one of the weaker ones (Figure 5.13a). The information transfer from AMV to itself is the strongest link in the entire system, something we already saw in 5.6. From AMV we see strong links to ENSO and TSD 5.13a, which are also the strongest incoming links to PDO 5.13c.

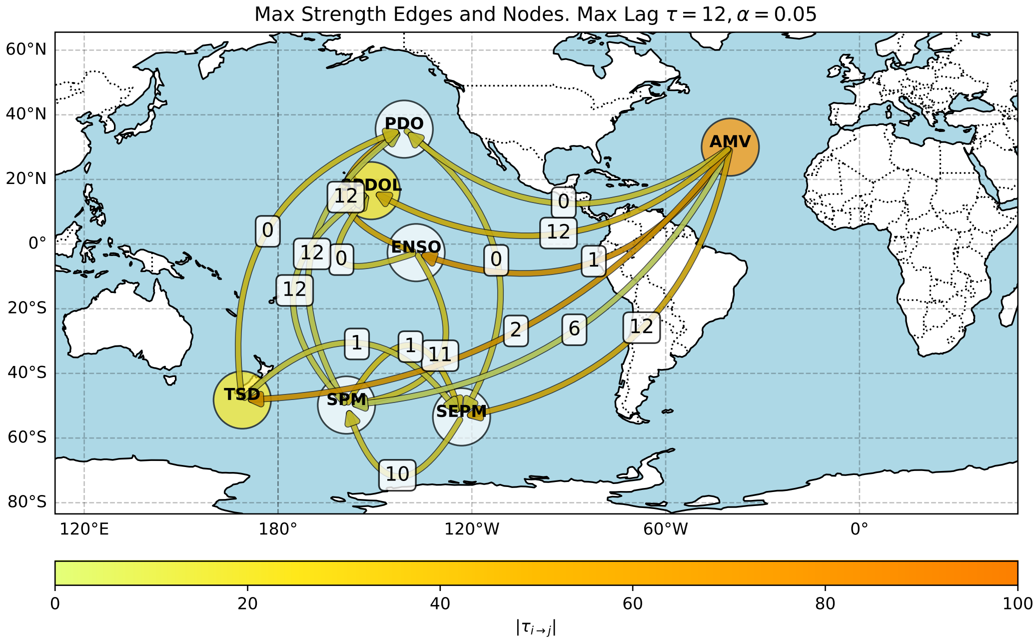
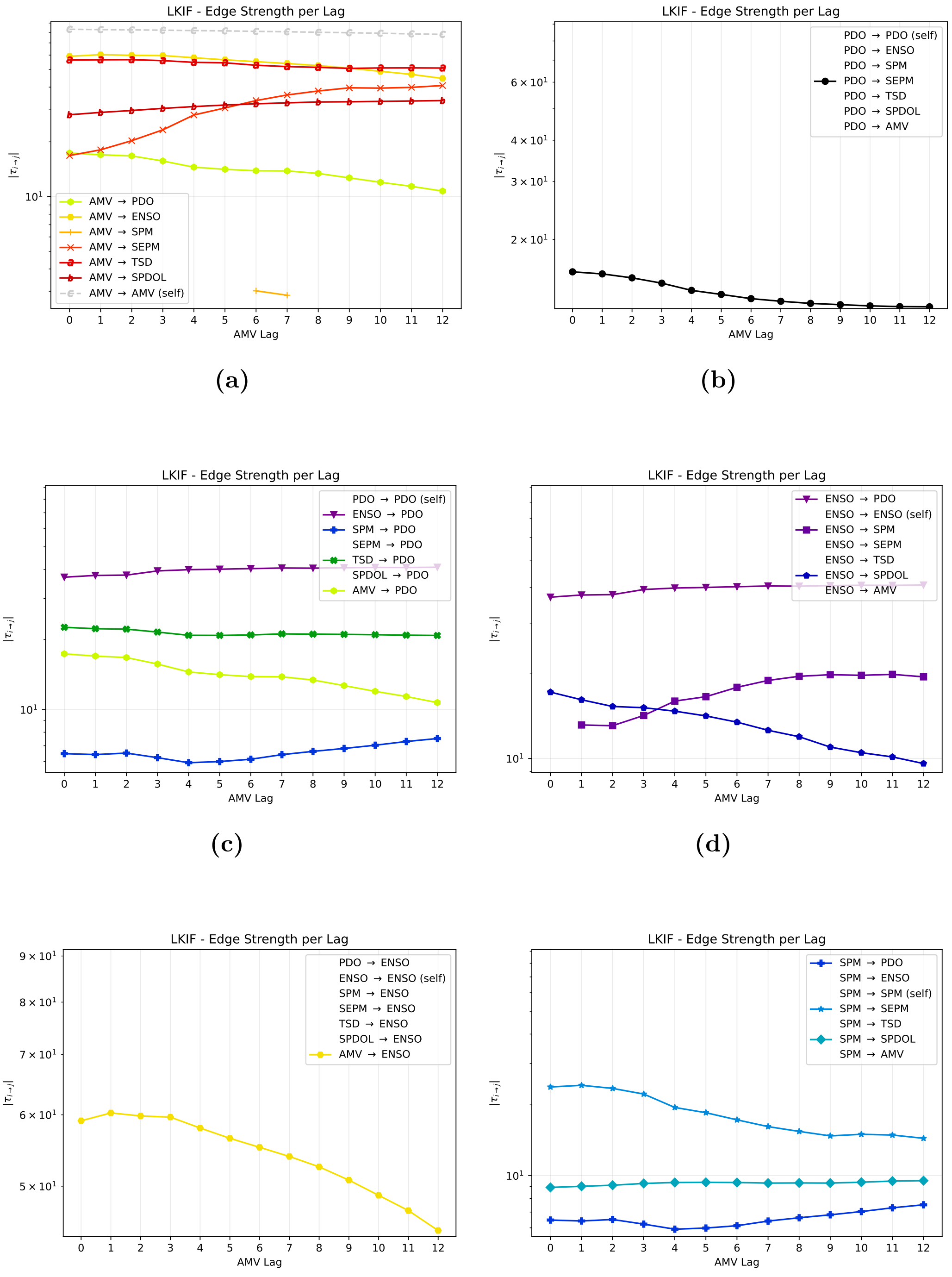


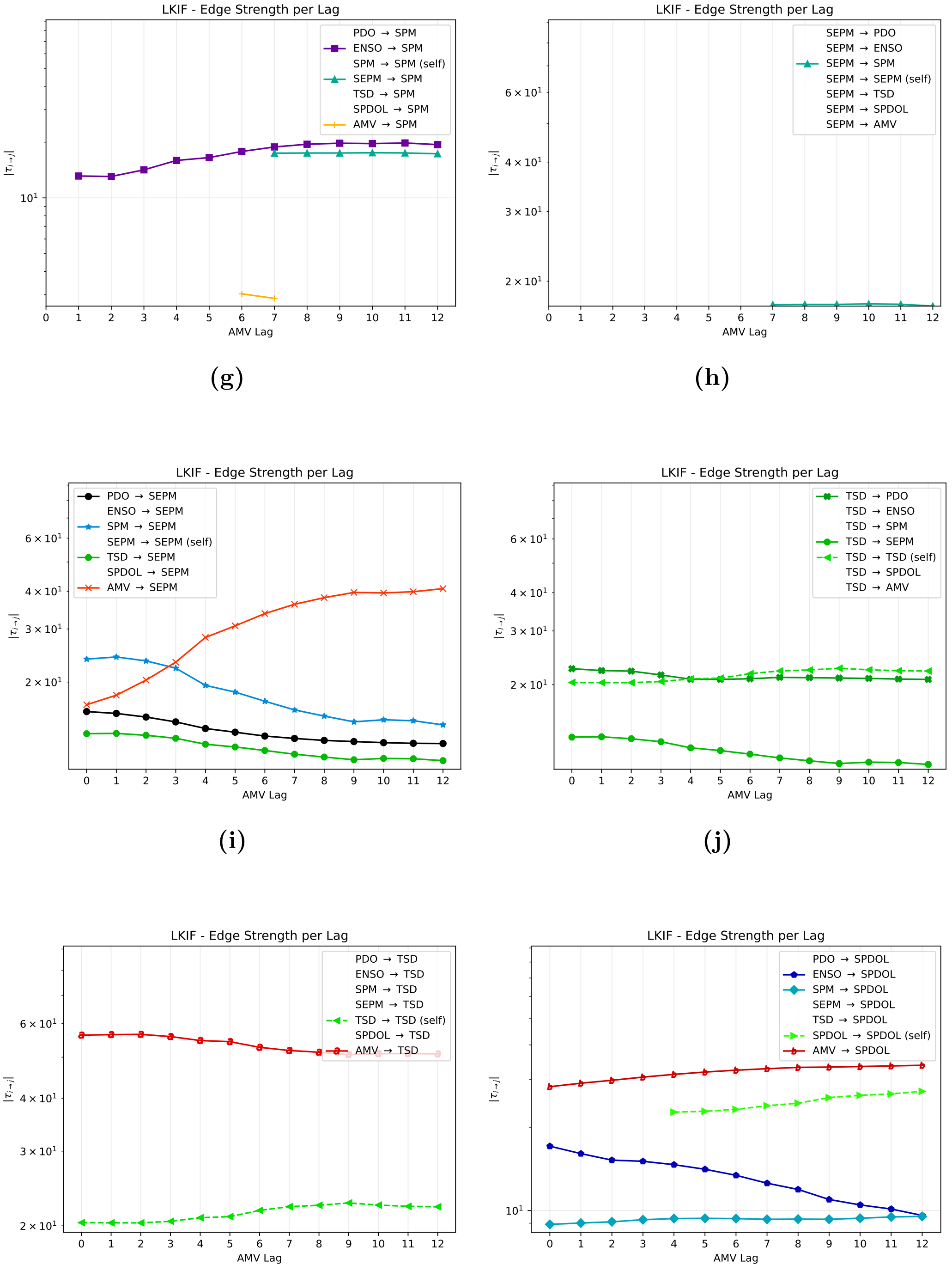
Figure 5.12 The strongest links found by LKIF when considering a system with AMV, PDO, and the first three varimax rotated principal components computed over 60S - 60N, 110E - 100W in the Pacific. The label on each link indicates the AMV lag at which the link is at its strongest. Edge colour indicates the absolute value of the normalised information flow. Node colour indicates the absolute value of the normalised information flow of a variable onto itself.


LKIF - Edge Strength per Lag
$|\tau_{i\to j}|$
10^1
AMV Lag
0 1 2 3 4 5 6 7 8 9 10 11 12
AMV → PDO
AMV → ENSO
AMV → SPM
AMV → SEPM
AMV → TSD
AMV → SPDOL
AMV → AMV (self)
LKIF - Edge Strength per Lag
6×10^1
4×10^1
3×10^1
2×10^1
PDO → PDO (self)
PDO → ENSO
PDO → SPM
PDO → SEPM
PDO → TSD
PDO → SPDOL
PDO → AMV
LKIF - Edge Strength per Lag
PDO → PDO (self)
ENSO → PDO
SPM → PDO
SEPM → PDO
TSD → PDO
SPDOL → PDO
AMV → PDO
LKIF - Edge Strength per Lag
ENSO → PDO
ENSO → ENSO (self)
ENSO → SPM
ENSO → SEPM
ENSO → TSD
ENSO → SPDOL
ENSO → AMV
LKIF - Edge Strength per Lag
9×10^1
8×10^1
7×10^1
6×10^1
5×10^1
PDO → ENSO
ENSO → ENSO (self)
SPM → ENSO
SEPM → ENSO
TSD → ENSO
SPDOL → ENSO
AMV → ENSO
LKIF - Edge Strength per Lag
SPM → PDO
SPM → ENSO
SPM → SPM (self)
SPM → SEPM
SPM → TSD
SPM → SPDOL
SPM → AMV

(a) (b)

(c) (d)

(e) (f)



(g) (h)

(i) (j)

(k) (l)

Figure 5.13 The edge strengths per AMV lag for LKIF. Each panel shows the strength at each lag for outgoing or incoming links for a single node in the graph.

Chapter 6

Discussion & Future Work

All three methods find causal links from the Atlantic to the Pacific Ocean, though PCMCI+ only does so when considering three varimax rotated principal components and the PDO. Granger causality and LKIF discover links from AMV to the Pacific more consistently than PCMCI+. Specifically, while Granger Causality and LKIF find direct, albeit weak, links from AMV to PDO, PCMCI+ never finds any. In the graphs found by LKIF , the AMV is always connected to all nodes in the Pacific, independently of the number of variables in the system. Across all three systems, PCMCI+ only finds a single connection from the AMV to the Pacific, namely when considering three varimax rotated components and the PDO, 5.4. This is different to the findings by Docquier et al. [34], where PCMCI+ found a connection from AMV to PDO at lag 8 and a weaker influence from PDO to AMV at lag 3. Their system included additional climate indices however, and it is difficult to account for this difference in systems when comparing these results. Furthermore, they focused on the 1950-2021 period and not 1871-2023 used in this thesis. Granger causality finds a lot more links than the other two methods. Recall that it operates on the principle of predictability between pairs of variables: if the past values of variable X improve the prediction of variable Y, X *Granger-causes* Y. However, this bivariate framework does not account for confounding factors or common causes - variables that may influence both X and Y simultaneously. As a result, Granger causality can mistakenly attribute causality to statistical associations that are actually due to shared influences from other, unmodeled variables. This leads to a higher number of causal links as seen in figures 5.2 and 5.8 many of which may actually be spurious. Therefore, the results should be taken with care. Krakovska et al. [78] provide a more detailed discussion of Granger Causality compared to other causal methods. PCMCI+, on the other hand, is designed to work in a multivariate framework, explicitly accounting for confounders, indirect links, and time-lagged dependencies across the entire system. Thanks to its conditioning set selection which identifies a minimal set of conditioning variables, it filters out spurious associations that arise from common causes or indirect paths, which Granger causality often overlooks. As a result, PCMCI+ tends to be more conservative, identifying fewer but ideally more robust and interpretable causal relationships. This selectivity is evident in the smaller number of links observed in figures 5.4 and 5.10, where PCMCI+ provides a sparser graph than Granger causality. The links found by LKIF agree with most of the previous literature on the direction of the connection from AMV to PDO (see 1.1.2), but differs from the structure uncovered by PCMCI+. The difference is likely due to the mechanics of the two methods, though it is difficult to understand the exact reason. Further investigation is needed to identify what causes this difference.

A common theme across the methods, when they find connections from the AMV to the Pacific, is that the strongest connection is not to PDO directly but to some other component. For example, when applied to a system containing the AMV, PDO, and first three rotated varimax components, LKIF finds a direct weak link from AMV to PDO (figure 5.7a) and stronger links to ENSO, SEPM, and SPDOL. ENSO then has a strong link to the PDO (figure 5.7e). Similarly, Granger Causality, when applied to the system with three varimax rotated components, AMV, and PDO, finds stronger links to SEPM and SPDOL than to the PDO (figure 5.3a) and of the incoming links to the PDO, the link from AMV is the weakest (figure 5.3d). This could hint at the fact that the influence from the AMV onto PDO is mediated by other regions, such as the tropical Pacific as described by Trascasa-Castro et al. [149] and Johnson et al. [66]. The timing of the connections may hint at the physical processes underlying interactions between climate variability modes. In particular, strong links at short time lags, such as the influence from the LKIF to the Pacific seen in Figure 5.6, may suggest the presence of an atmospheric teleconnection. Atmospheric teleconnections refer to large-scale, long-distance interactions between climate phenomena, where changes in one region can rapidly affect weather patterns in another [175]. A well-known example is ENSO, which influences climate across the Americas, Africa, and Asia (see Section 1). These atmospheric connections typically operate on timescales of days to weeks, aligning with the observed strong influences at lag 0 discovered by LKIF (figure 5.6), corresponding to a temporal resolution of less than one month. One possible mechanism for such rapid interaction is the Walker Circulation, which has been proposed by Kurcharski et al, Trascasa-Castro et al. and Levine et al. [79, 149, 83] as a pathway through which the AMV may affect the tropical Pacific. Other mechanisms through which the AMV can influence the Pacific are the Pacific North American (PNA) pattern [178], the Arctic Ocean [117], and the Polar Vortex [16], however none of these were modelled in any of our systems and so it is not possible to attribute the found links to these mechanisms. A consistently strong link for Granger causality and LKIF that is not present in the literature, is the edge from AVM to the Southern Pacific represented by the SEPM variable, in figure 5.3a and 5.7a. There are other modes in the South Pacific, such as the atmospheric Pacific South-American (PSA) pattern [63] which has been found to influence to the tropical North and South Atlantic [96]. However, it is important to note that the PSA pattern is atmospheric and that the links found by Meehl et al. [96] indicate an influence in the opposite direction (from the Pacific to the Atlantic) compared to our findings.

Subdividing the AMV into a tropical and subtropical component could be an interesting area to explore further. It should also be noted that in this work, the systems we compare only contain oceanic data in the form of SST. In contrast, the climate models (CM2.1, NCAR-CESM1, MIROC3.2m) used in previous research are fully coupled, meaning they include both oceanic and atmospheric components. The atmosphere plays a crucial role in driving, modulating, and responding to ocean variability. Hence, focusing solely on SST comes with some implications. Since oceanic variability is related to coupled ocean-atmosphere interactions, where atmospheric feedbacks can affect SST anomalies, focusing solely on SST, may miss important dynamical mechanisms or feedbacks that are essential to the full behaviour of the climate system.

Another aspect to consider is the linearity assumption of the three methods. All of them operate on the assumption that the relationships can be modelled with linear dynamics. Granger causality fits a linear regression, PCMCI+ uses a linear conditional independence test (PCMCI+ can be run with nonlinear CI tests like Gaussian Process Distance Correlation (GPDC) but we stuck with the linear Partial Correlation (ParCorr) to maintain a consistent approach for all three methods with regards to this aspect), and the MLE for LKIF is derived under a linear model assumption. Climate processes are often nonlinear however, and linear approximations therefore may not describe them accurately. Thus, it could be valuable to explore PCMCI+ with nonlinear CI tests as well as a nonlinear versions of Granger Causality and LKIF to see how they perform against the linear versions.

Chapter 7

Conclusion

This thesis aimed to investigate the causal interactions between variables in Atlantic and Pacific climate indices, using three distinct causal discovery methods: Granger Causality, PCMCI+, and LKIF. These methods were applied to smoothed versions of the AMV and PDO indices, as well as varimax rotated principal components in the Pacific, to uncover the underlying relationships and dynamics between these variables. All indices were computed from the ERSSTV5 dataset since the error of the AMV and PDO indices when compared to the ones by NOAA were the lowest. The results from the three methods revealed consistent findings that the Atlantic influences the Pacific, although the specifics of these interactions varied depending on the method and the number of variables included in the system. Granger Causality identified a dense network of connections, often linking the AMV to various Pacific components, including the PDO. However, its simplicity and bivariate nature led to a higher number of potentially spurious links, highlighting the need for more robust multivariate approaches. PCMCI+, designed to handle multivariate time series data, provided a more conservative and interpretable set of causal relationships. It identified fewer but potentially more robust links, particularly when considering three varimax rotated principal components alongside the AMV and PDO. This method's ability to account for confounding variables and indirect links made it a valuable tool for uncovering the underlying causal structure, although it struggled with smaller systems. LKIF, which measures the flow of information between variables, consistently found links from the AMV to the Pacific components, including the PDO. The strength of these links varied, with the strongest connections often being to components other than the PDO directly. This suggests that the influence of the AMV on the PDO may be mediated through other Pacific regions, such as the tropical Pacific, as proposed by previous research. The findings of this thesis align with much of the existing literature on the relationship between the AMV and PDO, reinforcing the notion that the Atlantic has a significant influence on Pacific climate variability. However, the study also highlighted the importance of considering the entire climate system, including other atmospheric variables, such as wind stress, which were not fully accounted for in this analysis. Future work could benefit from incorporating atmospheric data and exploring nonlinear causal discovery methods to better capture the complex dynamics of the climate system. In conclusion, this thesis contributes to the ongoing effort to understand the intricate relationships between the AMV and PDO, providing valuable insights into the causal interactions that govern these important climate indices. By leveraging advanced causal discovery methods, this work lays the groundwork for further research into the mechanisms driving global climate variability and the potential impacts of these interactions on future climate scenarios.

Bibliography

[1] Anteneh Z. Abiy et al. “Chapter 22 - Drought and climate teleconnection and drought monitoring”. In: *Extreme Hydrology and Climate Variability.* Ed. by Assefa M. Melesse, Wossenu Abtew, and Gabriel Senay. Elsevier, 2019, pp. 275–295. ISBN: 978-0-12-815998-9. DOI: `https://doi.org/10.1016/B978-0-12-815998-9.00022-1`. URL: `https://www.sciencedirect.com/science/article/pii/B9780128159989000221`.

[2] Emy Alerskans et al. “A transformer neural network for predicting near-surface temperature”. In: *Meteorological Applications* 29.5 (2022), e2098. DOI: `https://doi.org/10.1002/met.2098`. eprint: `https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/met.2098`. URL: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.2098`.

[3] Redazione ANSA. “Nel 2023 El Niño ha fatto impennare il livello dei mari”. In: *ANSA* (2024). URL: `https://www.ansa.it/canale_scienza/notizie/terra_poli/2024/03/24/nel-2023-el-nino-ha-fatto-impennare-il-livello-dei-mari_3e67d071-8cc9-42aa-ae06-4f20de766249.html`.

[4] Katinka Bellomo et al. “Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble”. In: *Climate Dynamics* 50 (May 2018). DOI: `10.1007/s00382-017-3834-3`.

[5] Jeroen Berrevoets et al. *Causal Deep Learning.* 2024. arXiv: `2303.02186 [cs.LG]`.

[6] Kaifeng Bi et al. “Accurate medium-range global weather forecasting with 3D neural networks”. In: *Nature* 619.7970 (2023), pp. 533–538. ISSN: 1476-4687. DOI: `10.1038/s41586-023-06185-3`. URL: `https://doi.org/10.1038/s41586-023-06185-3`.

[7] R. Black and E. Levenson. “Global Temperatures Set to Break Records, WMO Report Shows”. In: *The New York Times* (Mar. 2025). Accessed: 2025-03-20. URL: `https://www.nytimes.com/2025/03/18/climate/global-temperatures-wmo-report.html`.

[8] Niklas Boers et al. “Complex networks reveal global pattern of extreme-rainfall teleconnections”. In: *Nature* 566.7744 (Feb. 2019), pp. 373–377. ISSN: 1476-4687. DOI: `10.1038/s41586-018-0872-x`. URL: `https://doi.org/10.1038/s41586-018-0872-x`.

[9] Niklas Boers et al. “Prediction of extreme floods in the eastern Central Andes based on a complex networks approach”. In: *Nature Communications* 5.1 (Oct. 2014), p. 5199. ISSN: 2041-1723. DOI: `10.1038/ncomms6199`. URL: `https://doi.org/10.1038/ncomms6199`.

[10] Niklas Boers et al. “The South American rainfall dipole: A complex network analysis of extreme events”. In: *grl* 41.20 (Oct. 2014), pp. 7397–7405. DOI: `10.1002/2014GL061829`.

[11] Seth Borenstein. *2024: World’s Warmest Year on Record, Fueling Climate Disasters.* Accessed: 2025-03-20. Jan. 2025. URL: `https://apnews.com/article/climate-change-warming-hot-record-2024-disasters-12f899f071fcdbd051ad49a872611e92`.

[12] Florian Börgel et al. “Impact of the Atlantic Multidecadal Oscillation on Baltic Sea Variability”. In: *Geophysical Research Letters* 45.18 (2018), pp. 9880–9888. DOI: `https://doi.org/10.1029/2018GL078943`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL078943`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078943`.

[13] Jan Vom Brocke, René Riedl, and Pierre-Majorique Léger. “Application strategies for neuroscience in information systems design science research”. In: *Journal of Computer Information Systems* 53.3 (2013), pp. 1–13.

[14] Edward Bullmore and Olaf Sporns. “Complex brain networks: Graph theoretical analysis of structural and functional systems”. In: *Nature reviews. Neuroscience* 10 (Mar. 2009), pp. 186–98. DOI: `10.1038/nrn2575`.

[15] L.-P. Caron, M. Boudreault, and C. L. Bruyere. “Changes in large-scale controls of Atlantic tropical cyclone activity with the phases of the Atlantic multidecadal oscillation”. In: *Climate Dynamics* 44.7-8 (2015), pp. 1801–1821. DOI: `10.1007/s00382-014-2186-5`.

[16] Shangfeng Chen et al. “Atlantic multidecadal variability controls Arctic-ENSO connection”. In: *npj Climate and Atmospheric Science* 8.1 (2025), p. 44. ISSN: 2397-3722. DOI: `10.1038/s41612-025-00936-x`. URL: `https://doi.org/10.1038/s41612-025-00936-x`.

[17] Y. Chen et al. “Winter Eurasian cooling linked with the Atlantic multidecadal oscillation”. In: *Environmental Research Letters* 12.12 (2017), p. 125002. DOI: `10.1088/1748-9326/aa8de8`.

[18] Liao Chenfei et al. “The Impact of the Atlantic Multidecadal Oscillation on the Interannual Relationship Between the Storm Track and SST Over the North Atlantic in Spring”. In: *Atmosphere-Ocean* 59.3 (2021), pp. 152–164. DOI: `10.1080/07055900.2021.1927665`. eprint: `https://doi.org/10.1080/07055900.2021.1927665`. URL: `https://doi.org/10.1080/07055900.2021.1927665`.

[19] Amy Clement et al. “The Atlantic Multidecadal Oscillation without a role for ocean circulation”. In: *Science* 350.6258 (2015), pp. 320–324. DOI: `10.1126/science.aab3980`. eprint: `https://www.science.org/doi/pdf/10.1126/science.aab3980`. URL: `https://www.science.org/doi/abs/10.1126/science.aab3980`.

[20] University of Maine Climate Change Institute. *Sea Surface Temperature Daily Data*. Accessed: 2025-03-14. 2025. URL: `https://climatereanalyzer.org/clim/sst_daily/?dm_id=world2`.

[21] Copernicus Climate Change Service. *Global Climate Highlights 2024*. Accessed: 2025-03-20. 2025. URL: `https://climate.copernicus.eu/global-climate-highlights-2024`.

[22] Copernicus Climate Change Service. *Sea Surface Temperature - Climate Indicator*. Accessed: 2025-03-04. 2025. URL: `https://climate.copernicus.eu/climate-indicators/sea-surface-temperature`.

[23] Marc d’Orgeville and W. Richard Peltier. “On the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation: Might they be related?” In: *Geophysical Research Letters* 34.23 (2007). DOI: `https://doi.org/10.1029/2007GL031584`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL031584`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL031584`.

[24] Elizabeth M Daly and Mads Haahr. “Social network analysis for information flow in disconnected delay-tolerant MANETs”. In: *IEEE Transactions on Mobile Computing* 8.5 (2008), pp. 606–621.

[25] T. L. Delworth and M. E. Mann. “Observed and simulated multidecadal variability in the Northern Hemisphere”. In: *Climate Dynamics* 16.9 (2000), pp. 661–676. DOI: `10.1007/s003820000075`.

[26] Clara Deser and Adam S. Phillips. “Defining the Internal Component of Atlantic Multidecadal Variability in a Changing Climate”. In: *Geophysical Research Letters* 48.22 (2021). e2021GL095023 2021GL095023, e2021GL095023. DOI: `https://doi.org/10.1029/2021GL095023`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL095023`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL095023`.

[27] Clara Deser et al. “Sea Surface Temperature Variability: Patterns and Mechanisms”. In: *Annual Review of Marine Science* 2.Volume 2, 2010 (2010), pp. 115–143. ISSN: 1941-0611. DOI: `https://doi.org/10.1146/annurev-marine-120408-151453`. URL: `https://www.annualreviews.org/content/journals/10.1146/annurev-marine-120408-151453`.

[28] Nancy Devlin and Paul Hansen. “Health care spending and economic output: Granger causality”. In: *Applied Economics Letters* 8.8 (2001), pp. 561–564. DOI: `10.1080/13504850010017357`. eprint: `https://doi.org/10.1080/13504850010017357`. URL: `https://doi.org/10.1080/13504850010017357`.

[29] Alexander G Dimitrov, Aurel A Lazar, and Jonathan D Victor. “Information theory in neuroscience”. In: *Journal of computational neuroscience* 30 (2011), pp. 1–5.

[30] Thomas Dimpfl and Franziska Julia Peter. “Using transfer entropy to measure information flows between financial markets”. In: *Studies in Nonlinear Dynamics and Econometrics* 17.1 (2013), pp. 85–102.

[31] Mingzhou Ding, Yonghong Chen, and Steven L Bressler. “Granger causality: basic theory and application to neuroscience”. In: *Handbook of time series analysis: recent theoretical developments and applications* (2006), pp. 437–460.

[32] D. Docquier, S. Vannitsem, and A. Bellucci. “The rate of information transfer as a measure of ocean–atmosphere interactions”. In: *Earth System Dynamics* 14.3 (2023), pp. 577–591. DOI: 10.5194/esd-14-577-2023. URL: https://esd.copernicus.org/articles/14/577/2023/.

[33] D. Docquier, S. Vannitsem, and A. Bellucci. “The rate of information transfer as a measure of ocean–atmosphere interactions”. In: *Earth System Dynamics* 14.3 (2023), pp. 577–591. DOI: 10.5194/esd-14-577-2023. URL: https://esd.copernicus.org/articles/14/577/2023/.

[34] D. Docquier et al. “A comparison of two causal methods in the context of climate analyses”. In: *EGUsphere* 2023 (2023), pp. 1–31. DOI: 10.5194/egusphere-2023-2212. URL: https://egusphere.copernicus.org/preprints/2023/egusphere-2023-2212/.

[35] D. Docquier et al. “A comparison of two causal methods in the context of climate analyses”. In: *Nonlinear Processes in Geophysics* 31.1 (2024), pp. 115–136. DOI: 10.5194/npg-31-115-2024. URL: https://npg.copernicus.org/articles/31/115/2024/.

[36] David Docquier. *Codes to compute Liang index and correlation for comparison study.* Version v1.2. Sept. 2023. DOI: 10.5281/zenodo.8383534. URL: https://doi.org/10.5281/zenodo.8383534.

[37] Stillman Drake. *Galileo at Work: His Scientific Biography.* Facsimile edition. Mineola, NY: Dover Publications, 2003. ISBN: 9780486495422.

[38] C. T. Druery and William Bateson. “Experiments in Plant Hybridization”. In: *Journal of the Royal Horticultural Society* 26 (1901), pp. 1–32.

[39] Amitava Dutta. “Telecommunications and Economic Activity: An Analysis of Granger Causality”. In: *Journal of Management Information Systems* 17.4 (2001), pp. 71–95. DOI: 10.1080/07421222.2001.11045658. eprint: https://doi.org/10.1080/07421222.2001.11045658. URL: https://doi.org/10.1080/07421222.2001.11045658.

[40] ECMWF Copernicus Training. *Reanalysis and Climatology.* Online. Accessed: 2025-03-10. 2024. URL: https://ecmwf-projects.github.io/copernicus-training-c3s/reanalysis-climatology.html.

[41] James B Elsner. “Granger causality and Atlantic hurricanes”. In: *Tellus A: Dynamic Meteorology and Oceanography* 59.4 (2007), pp. 476–485.

[42] David B. Enfield, Alberto M. Mestas-Nuñez, and Paul J. Trimble. "The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S." In: *Geophysical Research Letters* 28.10 (2001), pp. 2077–2080. DOI: `https://doi.org/10.1029/2000GL012745`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2000GL012745`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000GL012745`.

[43] David B. Enfield, Alberto M. Mestas-Nuñez, and Paul J. Trimble. "The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S." In: *Geophysical Research Letters* 28.10 (2001), pp. 2077–2080. DOI: `https://doi.org/10.1029/2000GL012745`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2000GL012745`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000GL012745`.

[44] John Englart. *Warming may spike when Pacific Decadal Oscillation next switches to positive phase.* Accessed: 2024-12-08. Feb. 2014. URL: `https://takvera.blogspot.com/2014/01/warming-may-spike-when-pacific-decadal.html`.

[45] Stefan Feuerriegel et al. "Causal machine learning for predicting treatment outcomes". In: *Nature Medicine* 30.4 (Apr. 2024), pp. 958–968. ISSN: 1546-170X. DOI: `10.1038/s41591-024-02902-1`. URL: `https://doi.org/10.1038/s41591-024-02902-1`.

[46] Henry Fountain. "El Niño and La Niña: Understanding the Phenomena". In: *The New York Times* (2024). URL: `https://www.nytimes.com/2024/06/14/climate/el-nino-la-nina.html`.

[47] B. Fox-Kemper et al. "Ocean, Cryosphere and Sea Level Change". In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Ed. by V. Masson-Delmotte et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2021, pp. 1211–1362. DOI: `10.1017/9781009157896.011`.

[48] Andrew Freedman. "La Niña and Global Warmth: Coastal El Niño". In: *The Washington Post* (2025). URL: `https://www.washingtonpost.com/weather/2025/03/06/la-nina-global-warmth-coastal-el-nino/`.

[49] John R. Freeman. "Granger Causality and the Times Series Analysis of Political Relationships". In: *American Journal of Political Science* 27.2 (1983), pp. 327–358. ISSN: 00925853, 15405907. URL: `http://www.jstor.org/stable/2111021` (visited on 02/08/2025).

[50] Noah E Friedkin. "Information flow through strong and weak ties in intraorganizational social networks". In: *Social networks* 3.4 (1982), pp. 273–285.

[51] D. Giaquinto, W. Marzocchi, and J. Kurths. "Exploring meteorological droughts' spatial patterns across Europe through complex network theory". In: *Nonlinear Processes in Geophysics* 30.2 (2023), pp. 167–181. DOI: `10.5194/npg-30-167-2023`. URL: `https://npg.copernicus.org/articles/30/167/2023/`.

[52] Hans Gleisner. *Latitudinal Binning and Area-Weighted Averaging of Irregularly Distributed Radio Occultation Data.* Tech. rep. GRAS SAF Report 10. Accessed: 2025-03-10. Copenhagen, Denmark: Danish Meteorological Institute, Apr. 2011. URL: `http://www.grassaf.org/gsr_10`.

[53] C. W. J. Granger. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". In: *Econometrica* 37.3 (1969), pp. 424–438. ISSN: 00129682, 14680262. URL: `http://www.jstor.org/stable/1912791` (visited on 07/07/2024).

[54] A. Hannachi, I. T. Jolliffe, and D. B. Stephenson. "Empirical orthogonal functions and related techniques in atmospheric science: A review". In: *International Journal of Climatology* 27.9 (2007), pp. 1119–1152. DOI: `https://doi.org/10.1002/joc.1499`. eprint: `https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1499`. URL: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1499`.

[55] Abdel Hannachi. "A primer for EOF analysis of climate data". In: (Apr. 2004).

[56] Peter Hannam. "Global heating is making El Niño and La Niña forecasts less reliable, BoM says". In: *The Guardian* (2024). URL: `https://www.theguardian.com/australia-news/2024/sep/19/global-heating-is-making-el-nino-and-la-nina-forecasts-less-reliable-bom-says`.

[57] X. Hao, S. He, and H. Wang. "Asymmetry in the response of central Eurasian winter temperature to AMO". In: *Climate Dynamics* 47.7-8 (2016), pp. 2139–2154. DOI: `10.1007/s00382-015-2955-9`.

[58] Zeke Hausfather. "State of the Climate: 2024 Sets a New Record as the First Year Above 1.5°C". In: (Jan. 2025). Accessed: 2025-03-23. URL: `https://www.carbonbrief.org/state-of-the-climate-2024-sets-a-new-record-as-the-first-year-above-1-5c/`.

[59] Robin Hogg. "The Correct Way to Average the Globe". In: *Towards Data Science* (2019). Accessed: 2025-03-10. URL: `https://medium.com/towards-data-science/the-correct-way-to-average-the-globe-92ceecd172b7`.

[60] Yongmiao Hong, Yanhui Liu, and Shouyang Wang. "Granger causality in risk and detection of extreme risk spillover between financial markets". In: *Journal of Econometrics* 150 (Apr. 2007), pp. 271–287. DOI: `10.1016/j.jeconom.2008.12.013`.

[61] Boyin Huang et al. *NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5 [indicate subset used].* Accessed via NOAA/ESRL/PSD: `https://www.esrl.noaa.gov/psd/` [2025-04-07]. 2017. DOI: `10.7289/V5T72FNM`.

[62] Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Ed. by H.-O. Pörtner et al. Cambridge, UK and New York, NY, USA: Cambridge University Press, 2022, p. 3056. DOI: `10.1017/9781009325844`.

[63] Damien Irving and Ian Simmonds. “A New Method for Identifying the Pacific–South American Pattern and Its Influence on Regional Climate Variability”. In: *Journal of Climate* 29.17 (2016), pp. 6109–6125. DOI: `10.1175/JCLI-D-15-0843.1`. URL: `https://journals.ametsoc.org/view/journals/clim/29/17/jcli-d-15-0843.1.xml`.

[64] ISMAR CNR. *New Record Oceanic Temperatures.* Accessed: 2025-03-23. Jan. 2025. URL: `https://www.ismar.cnr.it/web-content/en/new-record-oceanic-temperatures/`.

[65] Japan Meteorological Agency. *Pacific Decadal Oscillation (PDO).* Accessed: 2025-03-23. 2024. URL: `https://ds.data.jma.go.jp/tcc/tcc/products/elnino/decadal/pdo.html`.

[66] Zachary Johnson et al. “Pacific Decadal Oscillation remotely forced by the equatorial Pacific and the Atlantic Oceans”. In: *Climate Dynamics* 55 (Aug. 2020). DOI: `10.1007/s00382-020-05295-2`.

[67] Jean Kaddour et al. *Causal Machine Learning: A Survey and Open Problems.* 2022. arXiv: `2206.15475`.

[68] Henry F. Kaiser. “The Varimax Criterion for Analytic Rotation in Factor Analysis”. In: *Psychometrika* 23.3 (1958), pp. 187–200. DOI: `10.1007/BF02289233`.

[69] S. Karmouche et al. “Regime-oriented causal model evaluation of Atlantic–Pacific teleconnections in CMIP6”. In: *Earth System Dynamics* 14.2 (2023), pp. 309–344. DOI: `10.5194/esd-14-309-2023`. URL: `https://esd.copernicus.org/articles/14/309/2023/`.

[70] M. T. Kayano and V. Capistrano. “How the Atlantic multidecadal oscillation (AMO) modifies the ENSO influence on the South American rainfall”. In: *International Journal of Climatology* 34.1 (2014), pp. 162–178. DOI: `10.1002/joc.3674`.

[71] Ryan Keisler. *Forecasting Global Weather with Graph Neural Networks.* 2022. arXiv: `2202.07575 [physics.ao-ph]`.

[72] Richard A. Kerr. “A North Atlantic Climate Pacemaker for the Centuries”. In: *Science* 288.5473 (2000), pp. 1984–1985. DOI: `10.1126/science.288.5473.1984`. eprint: `https://www.science.org/doi/pdf/10.1126/science.288.5473.1984`. URL: `https://www.science.org/doi/abs/10.1126/science.288.5473.1984`.

[73] Ji-Won Kim, Sang-Wook Yeh, and Eun-Chul Chang. “Combined effect of El Niño-Southern Oscillation and Pacific Decadal Oscillation on the East Asian winter monsoon”. In: *Climate Dynamics* 42.3 (2014), pp. 957–971. ISSN: 1432-0894. DOI: `10.1007/s00382-013-1730-z`. URL: `https://doi.org/10.1007/s00382-013-1730-z`.

[74] P. J. Klotzbach. “The influence of El Niño-Southern oscillation and the Atlantic multidecadal oscillation on Caribbean tropical cyclone activity”. In: *Journal of Climate* 24.3 (2011), pp. 721–731. DOI: `10.1175/2010JCLI3705.1`.

[75] P. J. Klotzbach and W. Gray. "Multidecadal variability in North Atlantic tropical cyclone activity". In: *Journal of Climate* 21.15 (2008), pp. 3929–3935. DOI: `10.1175/2008JCLI2162.1`.

[76] M. Knudsen et al. "Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years". In: *Nature Communications* 2 (2011), p. 178. DOI: `10.1038/ncomms1186`. URL: `https://doi.org/10.1038/ncomms1186`.

[77] Evan Kodra, Snigdhansu Chatterjee, and Auroop R Ganguly. "Exploring Granger causality between global average observed time series of carbon dioxide and temperature". In: *Theoretical and applied climatology* 104 (2011), pp. 325–335.

[78] Anna Krakovská et al. "Comparison of six methods for the detection of causality in a bivariate time series". In: *Phys. Rev. E* 97 (4 Apr. 2018), p. 042207. DOI: `10.1103/PhysRevE.97.042207`. URL: `https://link.aps.org/doi/10.1103/PhysRevE.97.042207`.

[79] Fred Kucharski et al. "Atlantic forcing of Pacific decadal variability". In: *Climate Dynamics* 46.7 (Apr. 2016), pp. 2337–2351. ISSN: 1432-0894. DOI: `10.1007/s00382-015-2705-z`. URL: `https://doi.org/10.1007/s00382-015-2705-z`.

[80] Remi Lam et al. "GraphCast: Learned Global Weather Forecasting". In: *Science* (2023). URL: `https://deepmind.com/research/publications/22598`.

[81] Minkyu Lee et al. "How Does Pacific Decadal Oscillation Affect Tropical Cyclone Activity Over Far East Asia?" In: *Geophysical Research Letters* 48.24 (2021). e2021GL096267 2021GL096267, e2021GL096267. DOI: `https://doi.org/10.1029/2021GL096267`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL096267`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL096267`.

[82] Suzanne S Lee. "Jumps and information flow in financial markets". In: *The Review of Financial Studies* 25.2 (2012), pp. 439–479.

[83] Aaron F. Z. Levine, Michael J. McPhaden, and Dargan M. W. Frierson. "The impact of the AMO on multidecadal ENSO variability". In: *Geophysical Research Letters* 44.8 (2017), pp. 3877–3886. DOI: `https://doi.org/10.1002/2017GL072524`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017GL072524`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL072524`.

[84] Peiyuan Li et al. "Regional Heatwave Prediction Using Graph Neural Network and Weather Station Data". In: *Geophysical Research Letters* 50 (Apr. 2023). DOI: `10.1029/2023GL103405`.

[85] X San Liang. "A study of the cross-scale causation and information flow in a stormy model mid-latitude atmosphere". In: *Entropy* 21.2 (2019), p. 149.

[86] X. San Liang. "Normalized Multivariate Time Series Causality Analysis and Causal Graph Reconstruction". In: *Entropy* 23.6 (2021). ISSN: 1099-4300. DOI: `10.3390/e23060679`. URL: `https://www.mdpi.com/1099-4300/23/6/679`.

[87] X. San Liang. “The Liang-Kleeman Information Flow: Theory and Applications”. In: *Entropy* 15.1 (2013), pp. 327–360. ISSN: 1099-4300. DOI: `10.3390/e15010327`. URL: `https://www.mdpi.com/1099-4300/15/1/327`.

[88] X. San Liang and Richard Kleeman. “Information Transfer between Dynamical System Components”. In: *Phys. Rev. Lett.* 95 (24 Dec. 2005), p. 244101. DOI: `10.1103/PhysRevLett.95.244101`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.95.244101`.

[89] Sindy Löwe et al. *Amortized Causal Discovery: Learning to Infer Causal Graphs from Time-Series Data.* 2022. arXiv: `2006.10833`.

[90] Pouya Manshour et al. “Causality and Information Transfer Between the Solar Wind and the Magnetosphere–Ionosphere System”. In: *Entropy* 23.4 (2021). ISSN: 1099-4300. DOI: `10.3390/e23040390`. URL: `https://www.mdpi.com/1099-4300/23/4/390`.

[91] Nathan Mantua and Steven Hare. “The Pacific Decadal Oscillation”. In: *Journal of Oceanography* 58 (Feb. 2002), pp. 35–44. DOI: `10.1023/A:1015820616384`.

[92] Nathan Mantua and Steven Hare. “The Pacific Decadal Oscillation”. In: *Journal of Oceanography* 58 (Feb. 2002), pp. 35–44. DOI: `10.1023/A:1015820616384`.

[93] Nathan J. Mantua et al. “A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production*”. In: *Bulletin of the American Meteorological Society* 78.6 (1997), pp. 1069–1080. DOI: `10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2`. URL: `https://journals.ametsoc.org/view/journals/bams/78/6/1520-0477_1997_078_1069_apicow_2_0_co_2.xml`.

[94] Robert Marschinski and Holger Kantz. “Analysing the information flow between financial time series: An improved estimator for transfer entropy”. In: *The European Physical Journal B-Condensed Matter and Complex Systems* 30 (2002), pp. 275–281.

[95] Marie C. McGraw and Elizabeth A. Barnes. “Memory Matters: A Case for Granger Causality in Climate Variability Studies”. In: *Journal of Climate* 31.8 (2018), pp. 3289–3300. DOI: `10.1175/JCLI-D-17-0334.1`. URL: `https://journals.ametsoc.org/view/journals/clim/31/8/jcli-d-17-0334.1.xml`.

[96] Gerald Meehl et al. “Atlantic and Pacific tropics connected by mutually interactive decadal-timescale processes”. In: *Nature Geoscience* 14 (Jan. 2021). DOI: `10.1038/s41561-020-00669-x`.

[97] Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. *Causal Transformer for Estimating Counterfactual Outcomes.* 2022. arXiv: `2204.07258`.

[98] Met Office. *Pacific Decadal Oscillation.* Met Office website. Accessed: 2025-04-05. 2024. URL: `https://weather.metoffice.gov.uk/learn-about/weather/oceans/pacific-decadal-oscillation`.

[99] Francesco Montagna et al. *Demystifying amortized causal discovery with transformers.* 2024. arXiv: `2405.16924`.

[100] Timothy J Mosedale et al. "Granger causality of coupled climate processes: Ocean feedback on the North Atlantic Oscillation". In: *Journal of climate* 19.7 (2006), pp. 1182–1194.

[101] NASA. *Global Temperature*. `https://climate.nasa.gov/vital-signs/global-temperature/?intent=121`. Accessed: 2025-04-30. 2025.

[102] NASA. *Pacific Decadal Oscillation (PDO)*. NASA Sea Level website. Accessed: 2025-04-05. 2024. URL: `https://sealevel.jpl.nasa.gov/data/el-nino-la-nina-watch-and-pdo/pacific-decadal-oscillation-pdo/`.

[103] National Center for Atmospheric Research. *Climate Indices*. Accessed: 2025-03-06. 2025. URL: `https://climatedataguide.ucar.edu/climate-data/indices`.

[104] National Oceanic and Atmospheric Administration (NOAA). *2024 Was World's Warmest Year on Record*. Accessed: 2025-03-20. 2025. URL: `https://www.noaa.gov/news/2024-was-worlds-warmest-year-on-record`.

[105] National Oceanic and Atmospheric Administration (NOAA). *El Niño and La Niña Impacts*. Accessed: 2025-03-06. 2024. URL: `https://www.noaa.gov/jetstream/enso_impacts`.

[106] National Oceanic and Atmospheric Administration (NOAA). *What are El Niño and La Niña?* Accessed: 2025-03-06. 2024. URL: `https://oceanservice.noaa.gov/facts/ninonina.html`.

[107] J. David Neelin et al. "ENSO theory". In: *Journal of Geophysical Research: Oceans* 103.C7 (1998), pp. 14261–14290. DOI: `https://doi.org/10.1029/97JC03424`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/97JC03424`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JC03424`.

[108] Matthew Newman et al. "The Pacific Decadal Oscillation, Revisited". In: *Journal of Climate* 29.12 (2016), pp. 4399–4427. DOI: `10.1175/JCLI-D-15-0508.1`. URL: `https://journals.ametsoc.org/view/journals/clim/29/12/jcli-d-15-0508.1.xml`.

[109] Eshaan Nichani, Alex Damian, and Jason D. Lee. *How Transformers Learn Causal Structure with Gradient Descent*. 2024. arXiv: `2402.14735`.

[110] NOAA. *What is the Atlantic Meridional Overturning Circulation (AMOC)?* National Ocean Service website. Accessed: 2025-04-05. 2024. URL: `https://oceanservice.noaa.gov/facts/amoc.html`.

[111] NOAA National Centers for Environmental Information. *Anomalies vs. Temperature - What's the Difference?* Accessed: 2025-03-09. 2025. URL: `https://www.ncei.noaa.gov/access/monitoring/dyk/anomalies-vs-temperature`.

[112] NOAA Physical Sciences Laboratory. *Atlantic Multidecadal Oscillation (AMO) Index*. `https://psl.noaa.gov/data/timeseries/AMO/`. Accessed: 2025-03-14.

[113] Peer Nowack et al. "Causal networks for climate model evaluation and constrained projections". eng. In: *Nature communications* 11.1 (Mar. 2020), p. 1415. ISSN: 2041-1723. DOI: `10.1038/s41467-020-15195-y`.

[114] Anne G. O'Carroll et al. "Observational Needs of Sea Surface Temperature". In: *Frontiers in Marine Science* 6 (2019). ISSN: 2296-7745. DOI: `10.3389/fmars.2019.00420`. URL: `https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2019.00420`.

[115] C. H. O'Reilly, M. Patterson, J. Robson, et al. "Challenges with interpreting the impact of Atlantic Multidecadal Variability using SST-restoring experiments". In: *npj Climate and Atmospheric Science* 6 (2023), p. 14. DOI: `10.1038/s41612-023-00335-0`. URL: `https://doi.org/10.1038/s41612-023-00335-0`.

[116] Floris van Ogtrop, Mukhtar Ahmad, and Carina Moeller. "Principal components of sea surface temperatures as predictors of seasonal rainfall in rainfed wheat growing areas of Pakistan". In: *Meteorological Applications* 21.2 (2014), pp. 431–443. DOI: `https://doi.org/10.1002/met.1429`. eprint: `https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/met.1429`. URL: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.1429`.

[117] Yuko M. Okumura et al. "North Pacific Climate Response to Freshwater Forcing in the Subarctic North Atlantic: Oceanic and Atmospheric Pathways". In: *Journal of Climate* 22.6 (2009), pp. 1424–1445. DOI: `10.1175/2008JCLI2511.1`. URL: `https://journals.ametsoc.org/view/journals/clim/22/6/2008jcli2511.1.xml`.

[118] A Jay Palmer, Christopher W Fairall, and WA Brewer. "Complexity in the atmosphere". In: *IEEE Transactions on Geoscience and Remote Sensing* 38.4 (2002), pp. 2056–2063.

[119] Jaime Palter. "The Role of the Gulf Stream in European Climate". In: *Annual review of marine science* 7 (Jan. 2015), pp. 113–37. DOI: `10.1146/annurev-marine-010814-015656`.

[120] Milan Paluš et al. "Causality, dynamical systems and the arrow of time". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.7 (July 2018), p. 075307. ISSN: 1054-1500. DOI: `10.1063/1.5019944`. eprint: `https://pubs.aip.org/aip/cha/article-pdf/doi/10.1063/1.5019944/13283420/075307\_1\_online.pdf`. URL: `https://doi.org/10.1063/1.5019944`.

[121] Mengting Pan et al. "Statistical calibrations to improve the 2–5-year prediction skill for SST over the North Atlantic". In: *Meteorology and Atmospheric Physics* 134 (June 2022). DOI: `10.1007/s00703-022-00888-4`.

[122] Christina Papagiannopoulou et al. "A non-linear Granger-causality framework to investigate climate–vegetation dynamics". In: *Geoscientific Model Development* 10.5 (2017), pp. 1945–1960.

[123] Marjorie Pett, Nancy Lackey, and John Sullivan. "Rotating the Factors". In: *Making Sense of Factor Analysis*. Thousand Oaks, California: SAGE Publications, Inc., 2003, pp. 131–166. DOI: `10.4135/9781412984898`. URL: `https://methods.sagepub.com/book/mono/making-sense-of-factor-analysis/toc`.

[124] John Pinneger and Caroline Rowland. *Turning Up the Heat: Advancing UK Science to Better Predict and Respond to Marine Heatwaves.* Accessed: 2025-03-14. 2024. URL: `https://marinescience.blog.gov.uk/2024/09/26/turning-up-the-heat-advancing-uk-science-to-better-predict-and-respond-to-marine-heatwaves/`.

[125] Mengjiao Qin et al. "Pacific Decadal Oscillation Forecasting With Spatiotemporal Embedding Network". In: *Geophysical Research Letters* 50.16 (2023). e2023GL103170 2023GL103170, e2023GL103170. DOI: `https://doi.org/10.1029/2023GL103170`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL103170`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2023GL103170`.

[126] Mikhail I Rabinovich et al. "Information flow dynamics in the brain". In: *Physics of life reviews* 9.1 (2012), pp. 51–73.

[127] S. Rahmstorf. "Thermohaline ocean circulation". In: *Encyclopedia of Quaternary Sciences.* Ed. by S. A. Elias. Elsevier, 2006, pp. 739–750.

[128] N. A. Rayner et al. "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century". In: *Journal of Geophysical Research: Atmospheres* 108.D14 (2003). Accessed via `https://www.metoffice.gov.uk/hadobs/hadisst/` [2025-04-07], p. 4407. DOI: `10.1029/2002JD002670`.

[129] Xianglin Ren et al. "The Pacific Decadal Oscillation modulated marine heatwaves in the Northeast Pacific during past decades". In: *Communications Earth & Environment* 4.1 (2023), p. 218. ISSN: 2662-4435. DOI: `10.1038/s43247-023-00863-w`. URL: `https://doi.org/10.1038/s43247-023-00863-w`.

[130] Alfredo Ruiz-Barradas, Sumant Nigam, and Argyro Kavvada. "The Atlantic Multidecadal Oscillation in twentieth century climate simulations: uneven progress from CMIP3 to CMIP5". In: *Climate Dynamics* 41.11 (2013), pp. 3301–3315. ISSN: 1432-0894. DOI: `10.1007/s00382-013-1810-0`. URL: `https://doi.org/10.1007/s00382-013-1810-0`.

[131] Jakob Runge. *Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets.* 2022. arXiv: `2003.03685 [stat.ME]`. URL: `https://arxiv.org/abs/2003.03685`.

[132] Jakob Runge and colleagues. *Tigramite: Causal Inference for Time Series.* `https://github.com/jakobrunge/tigramite`. Accessed: 2025-04-17. 2025.

[133] Jakob Runge et al. "Detecting and quantifying causal associations in large nonlinear time series datasets". In: *Science Advances* 5.11 (2019), eaau4996. DOI: `10.1126/sciadv.aau4996`. eprint: `https://www.science.org/doi/pdf/10.1126/sciadv.aau4996`. URL: `https://www.science.org/doi/abs/10.1126/sciadv.aau4996`.

[134] Jakob Runge et al. "Identifying causal gateways and mediators in complex spatio-temporal systems". In: *Nature Communications* 6.1 (2015), p. 8502. ISSN: 2041-1723. DOI: `10.1038/ncomms9502`. URL: `https://doi.org/10.1038/ncomms9502`.

[135] Yohan Ruprich-Robert et al. “Impacts of Atlantic multidecadal variability on the tropical Pacific: a multi-model study”. In: *npj Climate and Atmospheric Science* 4.1 (2021), p. 33. ISSN: 2397-3722. DOI: `10.1038/s41612-021-00188-5`. URL: `https://doi.org/10.1038/s41612-021-00188-5`.

[136] Skipper Seabold and Josef Perktold. “statsmodels: Econometric and statistical modeling with python”. In: *9th Python in Science Conference*. 2010.

[137] Copernicus Climate Change Service. *Sea Surface Temperature*. Accessed: 2025-03-14. 2025. URL: `https://climate.copernicus.eu/climate-indicators/sea-surface-temperature`.

[138] Anil K. Seth, Adam B. Barrett, and Lionel Barnett. “Granger Causality Analysis in Neuroscience and Neuroimaging”. In: *Journal of Neuroscience* 35.8 (2015), pp. 3293–3297. ISSN: 0270-6474. DOI: `10.1523/JNEUROSCI.4399-14.2015`. eprint: `https://www.jneurosci.org/content/35/8/3293.full.pdf`. URL: `https://www.jneurosci.org/content/35/8/3293`.

[139] Ali Shojaie and Emily B. Fox. “Granger Causality: A Review and Recent Advances”. In: *Annual Review of Statistics and Its Application* 9.Volume 9, 2022 (2022), pp. 289–319. ISSN: 2326-831X. DOI: `https://doi.org/10.1146/annurev-statistics-040120-010930`. URL: `https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-040120-010930`.

[140] Riccardo Silini et al. “Assessing causal dependencies in climatic indices”. In: *Climate Dynamics* 61.1 (2023), pp. 79–89. ISSN: 1432-0894. DOI: `10.1007/s00382-022-06562-0`. URL: `https://doi.org/10.1007/s00382-022-06562-0`.

[141] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. The MIT Press, Jan. 2001. ISBN: 9780262284158. DOI: `10.7551/mitpress/1754.001.0001`. URL: `https://doi.org/10.7551/mitpress/1754.001.0001`.

[142] Robert H. Stewart. *Introduction to Physical Oceanography*. Available in PDF format, licensed under CC BY-NC-SA. Texas A&M University, 2008. URL: `https://open.umn.edu/opentextbooks/textbooks/introduction-to-physical-oceanography`.

[143] Qiaohong Sun et al. “Century-scale causal relationships between global dry/wet conditions and the state of the Pacific and Atlantic Oceans”. In: *Geophysical Research Letters* 43.12 (2016), pp. 6528–6537. DOI: `https://doi.org/10.1002/2016GL069628`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL069628`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL069628`.

[144] Rowan T. Sutton and Daniel L. R. Hodson. “Atlantic Ocean Forcing of North American and European Summer Climate”. In: *Science* 309.5731 (2005), pp. 115–118. DOI: `10.1126/science.1109496`. eprint: `https://www.science.org/doi/pdf/10.1126/science.1109496`. URL: `https://www.science.org/doi/abs/10.1126/science.1109496`.

[145] Redaktion Tagesspiegel. “Unfassbare Temperaturen: 2023 war kein wärmstes Jahr wie jedes andere”. In: *Der Tagesspiegel* (2023). URL: `https://www.tagesspiegel.de/wissen/unfassbare-temperaturen-2023-war-kein-warmstes-jahr-wie-jedes-andere-11067010.html`.

[146] Alex Tank et al. “Neural Granger Causality”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1. ISSN: 1939-3539. DOI: `10.1109/tpami.2021.3065601`. URL: `http://dx.doi.org/10.1109/TPAMI.2021.3065601`.

[147] T. Tesch, S. Kollet, and J. Garcke. “Causal deep learning models for studying the Earth system”. In: *Geoscientific Model Development* 16.8 (2023), pp. 2149–2166. DOI: `10.5194/gmd-16-2149-2023`. URL: `https://gmd.copernicus.org/articles/16/2149/2023/`.

[148] Mingfang Ting et al. “Forced and Internal Twentieth-Century SST Trends in the North Atlantic”. In: *Journal of Climate* 22.6 (2009), pp. 1469–1481. DOI: `10.1175/2008JCLI2561.1`. URL: `https://journals.ametsoc.org/view/journals/clim/22/6/2008jcli2561.1.xml`.

[149] Paloma Trascasa-Castro et al. “Warm Phase of AMV Damps ENSO Through Weakened Thermocline Feedback”. In: *Geophysical Research Letters* 48.23 (2021). e2021GL096149 2021GL096149, e2021GL096149. DOI: `https://doi.org/10.1029/2021GL096149`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL096149`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021GL096149`.

[150] K. E. Trenberth and D. J. Shea. *AMO Index Data*. Provided by the Climate Analysis Section, NCAR, Boulder, USA. Updated yearly. Accessed DD Month YYYY. 2006. URL: `https://www.cgd.ucar.edu/cas/catalog/climind/AMO.html`.

[151] Kevin Trenberth, Rong Zhang, and National Center for Atmospheric Research Staff (Eds). *The Climate Data Guide: Atlantic Multi-decadal Oscillation (AMO) and Atlantic Multidecadal Variability (AMV)*. Last modified 2023-09-01. Retrieved from https://climatedataguide.ucar.edu/climate-data/atlantic-multi-decadal-oscillation-amo on 2024-11-13. 2023. URL: `https://climatedataguide.ucar.edu/climate-data/atlantic-multi-decadal-oscillation-amo`.

[152] Kevin E. Trenberth and John T. Fasullo. “An apparent hiatus in global warming?” In: *Earth’s Future* 1.1 (2013), pp. 19–32. DOI: `https://doi.org/10.1002/2013EF000165`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013EF000165`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013EF000165`.

[153] Kevin E. Trenberth and Dennis J. Shea. “Atlantic hurricanes and natural variability in 2005”. In: *Geophysical Research Letters* 33.12 (2006). DOI: `https://doi.org/10.1029/2006GL026894`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006GL026894`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026894`.

[154] Umberto Triacca. “Modes of climate variability and their relationships with interhemispheric temperature asymmetry: A Granger causality analysis”. In: *Theoretical and Applied Climatology* 143.3 (2021), pp. 1077–1081.

[155] Ulrich Tröhler. “Lind and scurvy: 1747 to 1795”. In: *Journal of the Royal Society of Medicine* 98.11 (Nov. 2005). Erratum in: J R Soc Med. 2006 Jan;99(1):45, pp. 519–522. DOI: `10.1177/014107680509801120`.

[156] U.S. Environmental Protection Agency. *Climate Change Indicators: Sea Surface Temperature.* Accessed: 2025-03-23. Jan. 2025. URL: `https://www.epa.gov/climate-indicators/climate-change-indicators-sea-surface-temperature`.

[157] S. Vannitsem and X.S. Liang. “Dynamical Dependencies at Monthly and Interannual Time Scales in the Climate System: Study of the North Pacific and Atlantic Regions”. In: *Tellus A: Dynamic Meteorology and Oceanography* 74.1 (2022), pp. 141–158. DOI: `10.16993/tellusa.44`. URL: `https://doi.org/10.16993/tellusa.44`.

[158] R. Vautard and M. Ghil. “Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series”. In: *Physica D: Nonlinear Phenomena* 35 (1989), pp. 395–424. DOI: `10.1016/0167-2789(89)90077-8`.

[159] Martin Vejmelka et al. “Non-random correlation structures and dimensionality reduction in multivariate climate data”. In: *Climate Dynamics* 44.9 (2015), pp. 2663–2682. ISSN: 1432-0894. DOI: `10.1007/s00382-014-2244-z`. URL: `https://doi.org/10.1007/s00382-014-2244-z`.

[160] Chunzai Wang et al. “El Niño and Southern Oscillation (ENSO): A Review”. In: *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment.* Ed. by Peter W. Glynn, Derek P. Manzello, and Ian C. Enochs. Dordrecht: Springer Netherlands, 2017, pp. 85–106. ISBN: 978-94-017-7499-4. DOI: `10.1007/978-94-017-7499-4_4`. URL: `https://doi.org/10.1007/978-94-017-7499-4_4`.

[161] Gang Wang et al. “The causality from solar irradiation to ocean heat content detected via multi-scale Liang–Kleeman information flow”. In: *Scientific Reports* 10.1 (2020), p. 17141. ISSN: 2045-2322. DOI: `10.1038/s41598-020-74331-2`. URL: `https://doi.org/10.1038/s41598-020-74331-2`.

[162] Shanshan Wang et al. “Combined effects of the Pacific Decadal Oscillation and El Niño-Southern Oscillation on Global Land Dry–Wet Changes”. In: *Scientific Reports* 4.1 (2014), p. 6651. ISSN: 2045-2322. DOI: `10.1038/srep06651`. URL: `https://doi.org/10.1038/srep06651`.

[163] Wenguang Wei, Zhongwei Yan, and Zhen Li. “Influence of Pacific Decadal Oscillation on global precipitation extremes”. In: *Environmental Research Letters* 16.4 (Mar. 2021), p. 044031. DOI: `10.1088/1748-9326/abed7c`. URL: `https://dx.doi.org/10.1088/1748-9326/abed7c`.

[164] Benjamin Werb and Daniel Rudnick. “Remarkable Changes in the Dominant Modes of North Pacific Sea Surface Temperature”. In: *Geophysical Research Letters* 50 (Feb. 2023). DOI: `10.1029/2022GL101078`.

[165] Seth Westra et al. “Interpreting variability in global SST data using independent component analysis and principal component analysis”. In: *International Journal of Climatology* 30.3 (2010), pp. 333–346. DOI: `https://doi.org/10.1002/joc.1888`. eprint: `https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1888`. URL: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1888`.

[166] Michael Wibral, Raul Vicente, and Joseph T Lizier. *Directed information measures in neuroscience.* Vol. 724. Springer, 2014.

[167] Frederik Wolf et al. “Event synchrony measures for functional climate network analysis: A case study on South American rainfall dynamics”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30.3 (Mar. 2020), p. 033102. ISSN: 1054-1500. DOI: `10.1063/1.5134012`. eprint: `https://pubs.aip.org/aip/cha/article-pdf/doi/10.1063/1.5134012/14623390/033102\_1\_online.pdf`. URL: `https://doi.org/10.1063/1.5134012`.

[168] World Climate Service. *What is a Climate Index?* Accessed: 2025-03-06. 2021. URL: `https://www.worldclimateservice.com/2021/09/13/climate-index/`.

[169] World Meteorological Organization. *Updated 30-year reference period reflects changing climate.* Accessed: 2025-04-16. May 2021. URL: `https://wmo.int/media/news/updated-30-year-reference-period-reflects-changing-climate`.

[170] World Meteorological Organization. *WMO Confirms 2024 Warmest Year on Record, About 1.55°C Above Pre-Industrial Level.* Accessed: 2025-03-20. 2025. URL: `https://wmo.int/news/media-centre/wmo-confirms-2024-warmest-year-record-about-155degc-above-pre-industrial-level`.

[171] World Meteorological Organization (WMO). *State of the Global Climate 2024.* English. WMO-No. 1368. Climate Statement. Geneva, 2025. URL: `https://library.wmo.int/doc_num.php?explnum_id=12034`.

[172] Fang Wu et al. “Information flow in social groups”. In: *Physica A: Statistical Mechanics and its Applications* 337.1-2 (2004), pp. 327–335.

[173] Sheng Wu et al. “Response of North Pacific and North Atlantic decadal variability to weak global warming”. In: *Advances in Climate Change Research* 9.2 (2018). Including special topic on mitigation for 1.5 °C: Scenarios and options, pp. 95–101. ISSN: 1674-9278. DOI: `https://doi.org/10.1016/j.accre.2018.03.001`. URL: `https://www.sciencedirect.com/science/article/pii/S1674927817301144`.

[174] Zhixiong Yao et al. “Predicting and Understanding the Pacific Decadal Oscillation Using Machine Learning”. In: *Remote Sensing* 16.13 (2024). ISSN: 2072-4292. DOI: `10.3390/rs16132261`. URL: `https://www.mdpi.com/2072-4292/16/13/2261`.

[175] Breanna Zavadoff and Marybeth Arcodia. *What are teleconnections? Connecting Earth's climate patterns via global information superhighways.* Accessed: 2025-05-07. Dec. 2022. URL: `https://www.climate.gov/news-features/blogs/enso/what-are-teleconnections-connecting-earths-climate-patterns-global`.

[176] R. Zhang and T. Delworth. "Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes". In: *Geophysical Research Letters* 33.17 (2006), p. L17712. DOI: `10.1029/2006GL026267`.

[177] R. Zhang and T. Delworth. "Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation". In: *Journal of Climate* 18.12 (2005), pp. 1853–1860. DOI: `10.1175/JCLI3460.1`.

[178] Rong Zhang and Thomas L. Delworth. "Impact of the Atlantic Multidecadal Oscillation on North Pacific climate variability". In: *Geophysical Research Letters* 34.23 (2007). DOI: `https://doi.org/10.1029/2007GL031601`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL031601`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL031601`.

[179] Rong Zhang, Thomas L. Delworth, and Isaac M. Held. "Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature?" In: *Geophysical Research Letters* 34.2 (2007). DOI: `https://doi.org/10.1029/2006GL028683`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006GL028683`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL028683`.

[180] W. Zhang et al. "Dominant role of Atlantic Multidecadal Oscillation in the recent decadal changes in western North Pacific tropical cyclone activity". In: *Geophysical Research Letters* 45.1 (2018), pp. 354–362. DOI: `10.1002/2017GL076397`.

[181] Yuan Zhang, John M. Wallace, and David S. Battisti. "ENSO-like Interdecadal Variability: 1900–93". In: *Journal of Climate* 10.5 (1997), pp. 1004–1020. DOI: `10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2`. URL: `https://journals.ametsoc.org/view/journals/clim/10/5/1520-0442_1997_010_1004_eliv_2.0.co_2.xml`.

Acknowledgements

I want to acknowledge and thank the people that have guided me during this thesis. First, my supervisor Paolo Tieri who wanted to explore the topic of climate networks and without whom I would never have started this work. Following, I want to thank Valerio Lembo from the Consiglio Nazionale delle Ricerche (CNR) and David Docquier from the Royal Meteorological Institute (RMI) of Belgium, for their expert guidance on the data and methods as well as their continuous feedback and suggestions on my work. Without them following me this closely, I would not have been able to produce this work. Finally, I would like to thank Alessio Bellucci from the CNR for adding valuable insight into the interpretation of the results.