DISSERTATION

CAUSAL INFERENCE USING OBSERVATIONAL DATA - CASE STUDIES IN CLIMATE SCIENCE

Submitted by

Savini M. Samarasinghe

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2020

Doctoral Committee:

Advisor: Dr. Imme Ebert-Uphoff

Dr. Chuck Anderson
Dr. Edwin Chong
Dr. Michael Kirby

ABSTRACT

CAUSAL INFERENCE USING OBSERVATIONAL DATA - CASE STUDIES IN CLIMATE SCIENCE

We are in an era where atmospheric science is data-rich in both observations (e.g., satellite/ sensor data) and model output. Our goal with causal discovery is to apply suitable data science approaches to climate data to make inferences about the cause-effect relationships between climate variables. In this research, we focus on using observational studies, an approach that does not rely on controlled experiments, to infer cause-effect. Due to reasons such as latent variables, these observational studies do not allow us to prove causal relationships. Nevertheless, they provide *data-driven hypotheses* of the interactions, which can enable us to get insights into the salient interactions as well as the timescales at which they occur.

Even though there are many different causal inference frameworks and methods that rely on observational studies, these approaches have not found widespread use within the climate or Earth science communities. To date, the most commonly used observational approaches include lagged correlation/regression analysis, as well as the bivariate Granger causality approach. We can attribute this lack of popularity to two main reasons. First is the inherent difficulty of inferring cause-effect in climate. Complex processes in the climate interact with each other at varying time spans. These interactions can be nonlinear, the distributions of relevant climate variables can be non-Gaussian, and the processes can be chaotic. A researcher interested in these causal inference problems has to face many challenges varying from identifying suitable variables, data, preprocessing and inference methods, as well as setting up the inference problem in a physically meaningful way. Also, the limited exposure and accessibility to modern causal inference approaches is another reason for their limited use within the climate science community.

In this dissertation, we present three case studies related to causal inference in climate science, namely, (1) causal relationships between the Arctic temperature and mid-latitude circulations, (2) relationships between the Madden Julian Oscillation (MJO) and the North Atlantic Oscillation (NAO) and (3) the causal relationships between atmospheric disturbances of different spatial scales (e.g., Planetary vs. Synoptic). We use methods based on probabilistic graphical models to infer cause-effect, specifically constraint-based structure learning methods, and graphical Granger methods. For each case study, we analyze and document the scientific thought process of setting up the problem, the challenges faced, and how we have dealt with the challenges. The challenges discussed include, but not limited to, method selection, variable representation, and data preparation. We also present a successful high-dimensional study of causal discovery in spectral space. The main objectives of this research are to make causal inference methods more accessible to a researcher/climate scientist who is at entry-level to spatiotemporal causality and to promote more modern causal inference methods to the climate science community. The case studies, covering a wide range of questions and challenges, are meant to act as a resourceful starting point to a researcher interested in tackling more general causal inference problems in climate.

ACKNOWLEDGEMENTS

First, I would like to convey my deepest gratitude to my research advisor, Dr. Imme Ebert-Uphoff, for her continuous support and guidance. Thank you for giving me the opportunity and freedom to grow and develop skills as a researcher. Thank you for your kindness and patience and for helping me navigate research and life as a graduate student through the last four and a half years.

My heartfelt thanks to my advising committee, Dr. Chuck Anderson, Dr. Edwin Chong and Dr. Michael Kirby for their constructive feedback, thoughtful comments, and encouragement. I would like to express my sincere thanks and appreciation to Dr. Elizabeth Barnes of the Department of Atmospheric Science, who has supported my work as a key research collaborator, advisor, and mentor. Thank you for providing me a platform to learn, discuss and appreciate the importance of data science applications in climate science by inducting me as part of the Barnes research group.

I am deeply indebted to all our climate science collaborators, Dr. Elizabeth Barnes, Dr. Yi Deng, Dr. Marie McGraw, Dr. Lantao Sun and Dr. Jason Furtado, without whom this research would not have been possible. Special thanks to Dr. Yi Deng of Georgia Tech for the immense support and contribution to the research related to scale-interaction processes contained in Chapter 5 of this dissertation.

I express my gratitude to the three educational institutions, Colorado State University, University of Moratuwa and Visakha Vidyalaya, the lecturers, and teachers who have taught me and shaped me as a researcher and as an individual. I gratefully acknowledge Dr. Clark Glymour and Dr. Joseph Ramsey of the Carnegie Mellon University for their willingness to share knowledge and for their advice related to causal inference methods. I would also like to extend my heartfelt thanks and sincere appreciation to Dr. Anura Jayasumana for his guidance and support throughout my journey as a graduate student.

I thank my friends and colleagues at Colorado State University, past and present, including Yajing Liu, Mahsa Ghorbani, Gunjan Mahindre, Marie McGraw, Sachintha Mendis, Shashika Mu-

ramudalige, Laksheen Mendis, Ben Toms, Kai-Chih Tseng, Lina Boljka, Gayatri Pendharkar, and Kirsten Mayer. Thank you for helping me better understand concepts in engineering, statistics, and atmospheric science. Thank you for encouraging me, being my support system and making my life as a graduate student memorable.

I appreciate all the love and support from Chandana DeSilva, Geetha Jayasumana and their families in helping me find a home away from *home*. I bare in my heart all my friends and family who have been my source of strength and motivation through this journey including my brother and sister-in-law Tharaka and Rasara Samarasinghe. Finally, I express my very profound gratitude to my parents, Chandana and Yamuna Samarasinghe, to my husband, Sajith Wijesuriya. Thank you for your unfailing love and support. This journey would have been impossible without you. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

This chapter explains the utility of observational studies for causal inference applications in climate science and motivates the research contained in this dissertation.

1.1 Causal Inference in Climate Science

The Earth is a complex system with many physical processes interacting across space and time. Understanding these interactions in a causal perspective, i.e., as cause and effect, can help us get a deeper understanding of the mechanisms governing Earth's climate. This type of understanding can help improve climate and weather prediction models by contributing to identify important predictors, timescales of interactions as well as pathways of interactions [1–10]. Another challenge associated with climate models is model uncertainty. Due to nonlinear dynamics, a large amount of parameterization (i.e., resolving sub-scale processes such as clouds, radiative transfer and microphysics), and model assumptions, different models simulating the same physical processes can provide very different results. In such cases, having an understanding of expected cause-effect relationships between climate variables can serve as a ground-truth or validation scheme to evaluate and select models that closely mimick the real-world [11–13]. Further, an understanding of causes and effects can help identify how the characteristics of atmospheric dynamics, as well as information flow, change as a response to climate change [14]. It can also allow to anticipate the effects of changing atmospheric and oceanic flow patterns that can change the frequencies of extreme weather events [15, 16]. These insights can be useful for decision making related to climate policy as well as disaster response.

A standard approach to identifying cause-effect in climate is through targeted climate modeling studies. As an example, a targeted modeling approach exploring connections between Arctic sea-ice concentrations and midlatitude circulations would involve using climate model runs to simulate the circulation responses to two scenarios related to sea-ice, i.e., for high and low sea ice concen-

trations. Provided that the simulations have been run for a sufficiently long time, the changes in midlatitude circulations in the two simulations can be attributed to the changes in sea-ice concentration. However, this type of study has a few limitations. First, the study only allows for exploring the causal interaction from sea-ice to circulations in isolation from the feedback from circulations to sea-ice [17] because the amount of sea ice is prescribed in the experimental setup. Second, the reliability of the inference depends on how well the dynamical model simulates the processes of interest. Furthermore, these simulations can be computationally expensive and time consuming. Setting up this type of experimental study can also be very difficult for certain complex dynamical processes.

In contrast to experimental studies, observational studies do not rely on specific controlled experiments. These studies use observed or simulated climate data *as is* and make inferences about the causal relationships using suitable data-driven approaches. Also, these methods do not ignore the feedbacks of the system, unlike targeted modeling approaches. They are thus allowing us to identify two-way causal interactions between variables. However, due to reasons such as latent variables, these methods do now allow us to prove causal relationships. They provide us with *data-driven hypotheses* of the potential causal interactions between the variables used in the study.

The climate scientist can use their background knowledge to validate whether the potential interactions identified by the inference approach are consistent with the physics/science governing the system and make a judgment on how to interpret the final results. Furthermore, these hypotheses can be useful to motivate and direct targeted modeling studies in the future. Chapter 2 reviews observational studies and causal inference methods based on observational studies that are relevant to this dissertation. The next section outlines the motivation for this research.

1.2 Problem Statement

There are many modern frameworks and approaches that rely on observational studies to infer causality. However, these methods are still not widely used in the climate or Earth science communities. To date, the most commonly used observational approaches include lagged correla-

tion/regression analysis as well as the bivariate Granger [18] causality approach. A main reason for this lack of popularity is that understanding the cause-effect relationships in the climate is an inherently challenging task. Many different complex processes in the climate interact with each other in varying time spans with some interactions happening every couple of days to others happening at annual to decadal time scales. These interactions can be nonlinear, the distributions of the relevant climate variables can be non-Gaussian, and the relevant processes can be chaotic.

A researcher interested in a specific science question is faced with the challenges of identifying the climatological subprocesses that are relevant to the question, determining how to extract the relevant signals from data, and how to set up the inference problem in a physically meaningful way. Also, for the inference to be useful, the method has be able to make correct inferences about the underlying system using sample data. Using appropriate preprocessing steps to amplify the causal signals in the presence of noise, and avoiding selection bias, measurement errors, and sampling errors can help the method capture a dependency structure that is *faithful* to the underlying probability distribution of the variables (see Section 2.2 for details).

The seasonal cycles and underlying trends of the climate variables need to be removed when the methods rely on stationarity assumptions. These trends can also act as hidden common causes among variables (where time can act as the hidden common cause). Standard causal inference methods assume that no hidden common causes are acting on two or more variables included in the study. This assumption is also known as the causal sufficiency assumption (see Section 2.2 for details). However, it is impossible to practically guarantee that this assumption has been met. There are causal inference methods such as the Fast Causal Inference (FCI - [19, 20]), time series FCI (tsFCI - [21]), Greedy FCI (GFCI - [22]) and Really Fast CI (RFCI - [23]) that in theory are supposed to identify asymptotically correct causal structures even in the presence of hidden common causes. However, in practice these methods are more computationally expensive, have limitations for time-series applications, and have not yet found wide-spread use.

In addition to the inherent difficulty of inferring causation in climate, the limited exposure and accessibility to modern causal inference approaches is another reason for their limited use within

the climate science community. These methods are currently popular among a small research community focused on developing inference methods [9] for applications such as neurosciences, bioinformatics, and econometrics, which rarely deal with spatiotemporal data. Further, the method descriptions and publicly available code implementations are often based on *jargon* specific to the causal inference community making them less accessible. Therefore, a researcher who is interested in using causal inference approaches for applications in climate will have to initially put effort to identify and understand suitable methods out of the many methods that are available, and customize them to spatiotemporal data. Also, as the correctness of some of these methods has been proved only at infinite sample size, their reliability at fixed sample sizes may vary requiring test cases and feasibility studies.

Furthermore, even after selecting a suitable causal inference approach, the researcher may have to make informed choices even within the numerous causal inference frameworks. For example, for the constraint-based and score-based methods the conditional independence test or scoring method needs to be selected based on the type of data, i.e., continuous or discrete or mixed, the type of relationships between variables, i.e., linear or non-linear, and distributional assumptions of the data, i.e., jointly Gaussian distributed, skewed, etc. It can be especially challenging to identify suitable inference methods that are scalable to high dimensional problems, particularly when the data is non-Gaussian and the relationships between the variables are non-linear.

Through this discussion, we see that the list of practical considerations and issues can be extensive. We have summarized the key points in Figure 1.1 below.

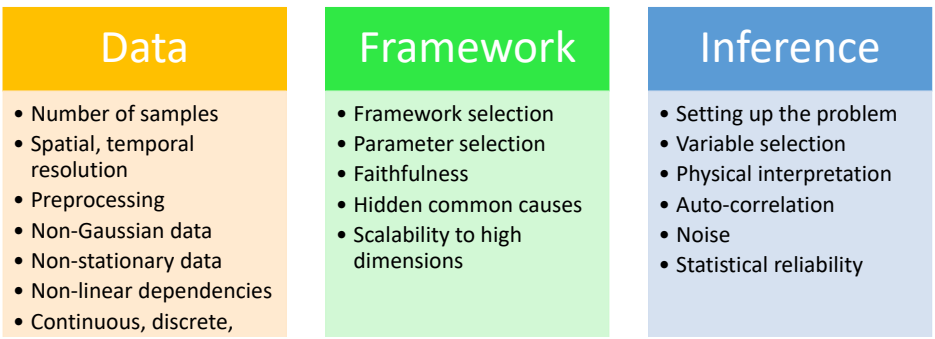


Figure 1.1: Summary of key considerations related to causal inference in climate science

1.3 Scope of the Disseration

The numerous considerations, questions, and challenges discussed in Section 1.2 are likely to impose a burden on a researcher interested in carrying out causal inference in the Earth sciences. This research attempts to ease this burden by investigating how several of these issues can be addressed. Specifically, we analyze a few causal inference questions in climate science, case-by-case, and document the scientific thought process of setting up the problem, the challenges faced, how the challenges are dealt with, and - in the process - also generate new scientific findings of interest to the climate science community. The main objective of this research is to make causal inference methods more accessible to a researcher/climate scientist who is at entry level to spatiotemporal causality.

The scope of this research is limited to constraint-based structure learning methods, and graphical Granger methods, which we hope will form the basis of the greater effort required to bridge the gaps between the causal inference and climate science communities. As case studies, we investigate (1) the causal relationships between the Arctic temperature and midlatitude circulations, (2) relationships between the Madden Julian Oscillation (MJO) and the North Atlantic Oscillation (NAO), and (3) the causal interactions of atmospheric disturbances at different spatial scales (e.g., Planetary vs. Synoptic). The key considerations relevant to each case study is listed in Table 1.1.

Table 1.1: The key considerations of each case study

Case study	Key considerations
1. Links between the Arctic and the midlatitude jet-stream	Method selection: PC vs. VAR-Granger and LASSO-Granger
2. Tropospheric and stratospheric causal pathways between the MJO and NAO	Problem setup Variable representation: MJO (cyclic phase) Mixed (continuous-discrete) variables Small sample size Utilizing conditional probabilities
3. Interactions between synoptic- and planetary-scale atmospheric disturbances	High dimensionality Variable representation: complex numbers Problem setup: from question to graph Extracting relevant information Validating results

As shown in Table 1.1, these case studies tackle a wide range of key considerations and challenges relevant to causal inference in climate. Therefore, we hope that the solutions and suggestions that we provide for each case study will act altogether as a resourceful starting point to many general causal inference problems in climate science. The additional contributions of this research are listed below.

Additional Contributions

1. Provide insights to atmospheric science. The case studies suggest the following.
    - Case 01: Arctic temperature is linked together with jet speed and position by two robust positive feedback loops that operate on timescales of 5-25 days.
    - Case 02: MJO can impact the NAO via both a tropospheric and stratospheric pathway. The stratosphere conditions the NAO to be conducive, or not, to MJO influence.
    - Case 03: The observational study provides insights into planetary-synoptic scale interactions that are consistent with the current understanding of the climate science community. a) Modulation of synoptic-scale disturbances by planetary-scale disturbances is best characterized by the flow of information from a zonal wavenumber-1 disturbance to a synoptic-scale circumglobal wave train whose amplitude peaks at the North Pacific and North Atlantic storm track region. b) The feedback of synoptic-scale to planetary-scale disturbances manifests itself as a zonal wavenumber-2 structure driven by synoptic-eddy momentum fluxes.
2. Provide step-by-step tutorials containing Python-based implementations of the bivariate Granger causality approach and the PC stable algorithm [24, 25]. The motivation is to make these methods more accessible to the climate science community.
3. First successful use of a causal inference algorithm in spherical harmonics space.

1.3.1 Organization of the Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 reviews the primary frameworks for causal inference of relevance to this dissertation. Chapters 3–5 cover the 3 case studies mentioned in the Table 1.1 in that respective order. These 3 chapters all contain research that have been published at peer-reviewed scientific journals [8, 26, 27] as listed below.

- S.M. Samarasinghe, M.C. McGraw, E. A. Barnes, and I. Ebert-Uphoff: A study of links between the Arctic and the midlatitude jet-stream using Granger and Pearl causality. *Environmetrics*, **30:e2540**, https://doi.org/10.1002/env.2540.

- E.A. Barnes, S.M. Samarasinghe, I. Ebert-Uphoff, and J. Furtado: Tropospheric and Stratospheric Causal Pathways between the MJO and NAO. *Journal of Geophysical Research: Atmospheres*, **124**, 9356-9371, https://doi.org/10.1029/2019JD031024.

- S. M. Samarasinghe, Y. Deng, and I. Ebert-Uphoff: A Causality-Based View of the Interaction between Synoptic- and Planetary-Scale Atmospheric Disturbances. *Journal of the Atmospheric Sciences*, **77**, 925-941, https://doi.org/10.1175/JAS-D-18-0163.1.

Chapter 6 provides conclusions and discusses suggestions for future work.

2 Observational Studies for Causal Inference

The formal framework of causality that allows to positively prove or disprove causal relationships is based on *interventions* [20, 28, 29]. In an intervention analysis, a scientist intervenes or performs experiments in the system of interest by changing/manipulating the state of any variable, and then observes the cascading effects of that change on other variables. However, performing intervention analysis in climate science is difficult, and requires targeted modeling approaches that use dynamical models and specific experimental design. As discussed in Chapter 1, these modeling approaches have certain limitations.

In such situations, resorting to an observational study (instead of an intervention analysis) that uses observed or simulated data (as is without any experimentation) can help get insights into questions that can otherwise be difficult to answer. Due to reasons such as the potential existence of latent variables, we cannot prove any interactions through an observational study. However, we can nevertheless disprove causal interactions through an observational study (see Section 2.2.1 for details). Thus, an observational study only yields *hypotheses of potential causal interactions* that need to be further evaluated by a climate scientist.

Even though observational studies cannot be used to prove causal relationships, causal discovery methods based on observational studies have been proven to be useful to gain valuable insights from observational data in many domains such as social sciences, economics, biology, and bioinformatics when interventions may be impossible, difficult or unethical [20, 30]. Understanding interactions between different regions of the brain using fMRI data [31] and identifying gene regulatory networks using gene expression data [32, 33] are some examples. Over the years, many different methods and efficient algorithms have been developed to identify potential cause-effect relationships from observational studies. However, most of these methods are still not very well known within the climate science community. The common approaches to identify associative and causal relationships in climate science to-date are based on correlation analysis, simple/traditional

lagged regression, and to some extent on bivariate Granger causality [18, 34] approach. Equation 2.1 provides an example of predicting a time series $Y = \{y_t\}_{t=1}^{T}$ using the past of $X = \{x_t\}_{t=1}^{T}$ using a simple lagged regression approach. Note that the auto-correlation/memory of Y is being ignored in this model.

$$\text{Simple lagged regression:} \quad y_t = \beta + a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_{t-p} + e, \quad t = (p+1), \ldots, T \tag{2.1}$$

Here a_i (for integer $i \in [1, p]$) are scalar regression coefficients and β is the model intercept. p is the maximum lag of X used for prediction. e is the error term of the model.

Even though correlation analysis and lagged regression methods are fundamental tools to the field of climate science, there are limitations of using them for causal analysis. For example, a simple lagged regression model predicting Y using the past of X (but excluding the past of Y as in Equation 2.1) might give the impression that X causes Y, when in reality (a) Y causes X but the analysis is affected by the autocorrelation of Y (see [35] or [36] for details), (b) X and Y are connected indirectly through another variable Z, (c) X and Y do not have any causal interactions between them but are driven by a common variable Z (Z is a common cause of X and Y). Figure 2.1 shows these three examples graphically.

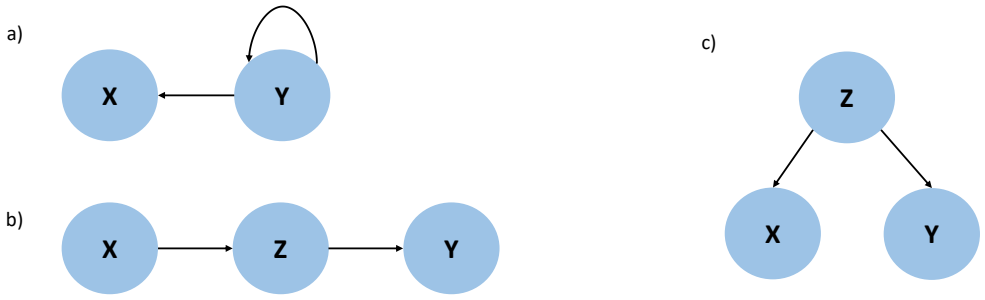


Figure 2.1: Sample scenarios where simple lagged regression could incorrectly indicate that X causes Y. Figure adapted from [4].

Causal discovery frameworks that are based on rigorous definitions have supported the development of methods that seek to correctly determine the directionality of causation, distinguish

between direct and indirect connections, handle the memory and feedbacks in a system, and in certain cases be able to identify causal connections correctly even in the presence of hidden (latent) common causes in contrast to correlation analysis and simple lagged regression. Out of the different causal inference approaches, we list the two key approaches below.

- **Granger causality** (e.g., VAR-Granger, LASSO-Granger, Transfer entropy, and other nonlinear Granger approaches. See Section 2.1 for details.)

- **Graphical causal model approaches** (e.g., constraint-based structure learning methods such as PC [37], PC-Stable [38] and Fast Causal Inference (FCI) [19, 20], score based structure learning methods such as Fast-GES (FGES) [39], Fast Adjacency Skewness (FASK), Greedy FCI (GFCI) [22], and Gaussian graphical methods. See Section 2.2 for details.)

These methods based on observational studies also provide advantages over the traditional targeted modeling approach by allowing to identify causal pathways relative to other drivers and feedbacks and by allowing for a direct comparison of results from observations and models.

Despite these advantages, only a handful of research groups have used more modern causal inference approaches for climate science applications. [1,36,40–47] are a few examples. This lack of popularity may be attributed to the fact that inferring causality in climate science is inherently more difficult compared to other disciplines, as discussed in Section 1.2. In addition, Runge et al. [9] provide an excellent overview of causal inference in Earth system sciences. Amidst challenges, using causal discovery approaches to identify causal interactions provides a major step forward from the traditional approaches. Scalable implementations of these methods allow us to tackle high dimensional problems and infer causality among a large number of climate variables. The graphical causal model approaches additionally provide a visual representation of the causal structure that eases the interpretation compared to the pairwise comparisons of the bivariate Granger analysis. Section 2.1 provides an overview of Granger causality and Section 2.2 provides and overview of graphical causal models.

2.1 Granger Causality

Granger causality [18, 34] is a statistical causal inference approach first introduced to econometrics in the 1960s. In bivariate Granger causality, a time series X is said to Granger cause another time series Y if the past of X contains unique information that is relevant to the current state of Y, that is not contained in the past of Y or all other past information *in the universe*. To determine whether X Granger causes Y in a practical and operational setting, we can ask a simplified question based on the predictability of these variables – 'does the past of X help to predict the current state of Y beyond Y's ability to predict itself?' An alternative approach would be to pose this question based on transfer entropy (an information-theoretic measure) and ask 'does X reduce the uncertainty of the current state of Y, beyond the degree to which the past of Y can already reduce the uncertainty of Y?' [48]

Granger causality is a well-established, practical, and operational definition of causality for time series variables and is the best known causal inference framework in climate science to date. For an introductory overview and illustrations of the utility of the Granger causality approach over simple lagged regression analysis, see [36]. This approach has been used and found useful in climate attribution studies, for example, to understand the effect of CO_2 on global temperature [49,50], to examine the impact of sea surface temperatures on the North Atlantic Oscillation (NAO) [51, 52] or the Arctic on midlatitude circulation [36].

2.1.1 Linear Bivariate Granger Causality

A common approach to identify Granger causality is based on the question of predictability posed before and involves statistically comparing the error performance of two prediction models. Namely, a *restricted model* where Y is predicted using the past values of Y alone, and an *unrestricted model* where Y is predicted using the past values of both X and Y. Then a statistical test is carried out to determine whether the unrestricted model significantly improves the predictability of Y compared to the restricted model. If so, X is said to Granger cause Y.

Equations 2.2 and 2.3 show examples of restricted and unrestricted linear auto-regression models predicting Y.

$$\text{Restricted:} \quad y_t = \gamma + b_1 y_{t-1} + b_2 y_{t-2} + \ldots + b_s y_{t-s} + e_r \tag{2.2}$$

$$\text{Unrestricted:} \quad y_t = \beta + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_s y_{t-s} + c_1 x_{t-1} + c_2 x_{t-2} + \ldots + c_p x_{t-p} + e_u \tag{2.3}$$

Here a_i, b_i (for integer $i \in [1, s]$) and c_j (for integer $j \in [1, p]$) are scalar regression coefficients and β and γ are model intercepts. s is the maximum lag of Y and p is the maximum lag of X used for prediction. e_r and e_u are the error terms of the restricted and unrestricted models respectively. To identify the suitable number of lags, we can use information criteria such as the Akaike information criterion (AIC) or the Bayesian information (BIC) or use an appropriate cross-validation scheme [53].

In Equation 2.3, X is not a Granger cause of Y if all the regression coefficients related to X, i.e., $c_1, \ldots, c_p$, are equal (or very close) to zero. In other words, this is when X does not significantly contribute to the predictability of Y. Thus the null-hypothesis for non-causality is $H_0 : c_1 = c_2 = \ldots = c_p = 0$, and this can be tested for example by comparing the variances of the residuals of the restricted and unrestricted models (see for example, [48]). We provide a step-by-step tutorial containing a Python-based implementation of the bivariate Granger causality approach in the Data Analysis Tools for the Atmospheric Sciences (DATAS) gateway [25].

2.1.2 Beyond Linear Bivariate Granger Causality Methods

The bivariate Granger causality approach will neither allow us to distinguish between direct vs. indirect relationships nor identify common drivers in the system. However, setting up the problem in the multivariate setting can help overcome these problems. Section 3.4.1 shows how Granger causality approaches based on Vector Auto Regression (VAR) as well as regularized regression can be useful to identify Granger causes in a multivariate setting. The relationships identified using these approaches can be represented in a graphical way as graphical Granger models (see [54] for example).

In place of the linear auto-regressive models used in this discussion, we can use appropriate nonlinear models to identify nonlinear dynamics between variables. For example, [55] uses a nonlinear Granger causality approach based on random forests while [56, 57] use models based on Artificial Neural Networks. Finally, we stress again that any approach based on observational data only identifies potential causal interactions and does not guarantee 'true' causality. The insights inferred from this type of study will be in the context of the variables used and subject to the assumptions used by the inference methods such as stationarity of variables.

2.2 Graphical Causal Models

For a set of random variables, the conditional independence structure between them can be represented as a probabilistic graphical model in which the variables of interest are represented using graph vertices/nodes, and the direct probabilistic dependencies between them are represented using graph edges. The directness of a dependency is always conditional (or in context) to the variables included in the model. We can use a Bayesian network to represent a directed structure as a Directed Acyclic Graph (DAG) or a Markov random field to represent an undirected structure. In the context of our research, we will only focus on directed structures.

A directed graph $G = (V, E)$ contains a set of vertices $V = \{V_1, V_2, \ldots, V_n\}$ and a set of directed edges E that connect pairs of vertices. For a directed edge between two vertices $V_i \rightarrow V_j$, we say that the vertex V_i is the parent of the child V_j. A DAG is a directed graph that has no cyclic paths – a path is cyclic when it starts from a vertex and then cycles back to the same vertex, for example, $V_i \rightarrow V_j \rightarrow V_i$. A DAG can be used to represent the conditional independence relationships of a probability distribution, i.e., as a Bayesian network (for details on Bayesian networks, see for example [58]). Figure 2.2 shows an example related to three discrete random variables relating to how a specific patient's tendency to get an asthma attack can be connected to the occurrence of forest fires and the air quality in the patient's neighborhood.

1. Forest fire - has states yes/no based on whether there is a forest fire in the neighborhood.
2. Air quality - has states good/bad based on air quality indices in the neighborhood.
3. Asthma attack - has states yes/no based on whether a specific asthma patient gets an attack.

Figure 2.2: Graph representing relationships between forest fire, air quality and tendency to get an asthma attack for a specific patient.

As shown in Figure 2.2, whether there is a forest fire or not can have a direct impact on the air quality of a geographical region. Having a forest fire can lead to poor air quality, which can increase the tendency of asthma attacks. Based on this graph, there is a *direct* dependency between forest fire and air quality whilst the dependency between forest fire and asthma attack is an *indirect* one. If there is a forest fire, the patient will be more likely to get an asthma attack. However, if we already know about the air quality of the region, knowing whether there is a forest fire or not will not give us additional information about the patient's tendency to get an attack. This means that the air quality variable blocks the information flow between the forest fire and asthma attack variables. In probability terms, this means that forest fire and asthma attacks are conditionally independent, given the air quality variable.

Pearl [28, 29] and Spirtes et al. [20] pioneered the research in using DAGs in causal inference by coming up with rigorous mathematical frameworks that allow to systematically relate DAGs to probability distributions, arriving at conditions that are required to interpret these graphs as causal models, and introducing methods that allow learning these structures from data efficiently. We need to make the following assumptions about the underlying distributions to be able to interpret a resulting Bayesian network as a graphical causal model [20, 59].

1. The model is causally sufficient.

   This means that there are no hidden common causes, aka latent confounders, to the variables included in the model. If any two nodes X and Y of the graph have a common cause L in the underlying system, then L must also be included in the graph. Figure 2.3 shows an example of a hidden common cause.

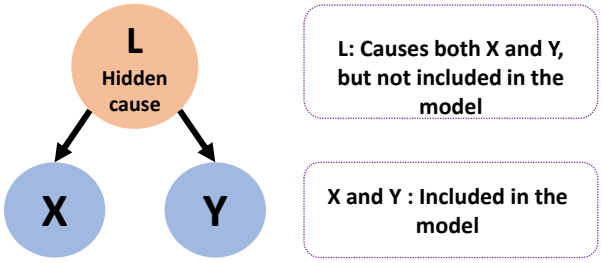


Figure 2.3: L is a hidden common cause to the variables X and Y in the model.

It is important to note that based on this assumption, the graph needs to be *complete* with respect to common causes. However, this assumption does not imply that the graph cannot have unobserved variables. For example, there can be an unobserved variable acting as an indirect cause or an unobserved variable driving only a single variable in the model [59].

2. The model satisfies the Causal Markov condition

   The Causal Markov condition states that each variable X in the graph, given its direct causes (i.e., its parents), is independent of every other variable that is neither a direct cause of X nor a descendant (an effect) of X [20,59]. This property allows to represent the joint distribution among all the variables of the graph in a minimal fashion, as in Equation 2.4 below.

$$P(V_1, V_2, \ldots, V_n) = \prod_{i=1}^{n} P(V_i | parents(V_i)). \tag{2.4}$$

3. The model is Faithful

   This means that the graph is a good representation of the underlying probability distribution. If a conditional independence does not hold in the probability distribution, then it should not hold in the graph either.

When these conditions are met, we can interpret a direct connection from node X to node Y, i.e., $X \rightarrow Y$, as X is a direct cause of Y relative to the variables included in the model. Unfortunately, the required conditions to interpret a DAG as a causal model can be violated in practical settings, especially the causal sufficiency condition. Causal sufficiency can be violated

because we usually have to make an educated decision about the variables that are relevant to the model, resulting in a risk of missing common causes. In addition, some variables that cannot be directly measured may need to be represented through proxies. This violation of conditions explains why interactions identified by a causal discovery algorithm need to be treated with caution as potential interactions. The interactions are either true causal interactions or spurious interactions that are due to reasons such as hidden common causes or inadequacies of faithfulness. The next section explains the primary algorithm used in this research to derive a graphical causal model from data. In addition, see [1] for a detailed discussion of graphical causal models and constraint-based structure learning for applications in climate science.

2.2.1 Constraint–Based Structure Learning

Constraint-based structure learning methods provide efficient approaches to learn DAGs that encapsulate the conditional independencies that are seen in observational data. As mentioned before, an observational analysis can only give information about potential causal relationships - we cannot use such a study to prove the *sufficient* conditions of causality. Nevertheless, we can still eliminate or disprove causal relationships by using the *necessary* conditions, i.e., we can say that two variables are not causally interacting if the relationship between them does not at least satisfy the necessary conditions. With constraint-based structure learning, we use the necessary condition that two variables X and Y are conditionally dependent given all other variables if there is a direct causal relationship between them. However, the necessary condition gets violated if we find that X and Y are conditionally independent given a set of variables (excluding X and Y), disproving the causal relationship.

As shown in Figure 2.4, the structure learning methods start with a fully connected undirected graph (Figure 2.4, blocks 1 and 2). The initial assumption is that all pairs of variables have cause-effect connections to each other. Then suitable conditional independence tests are used to test whether each pair of variables are conditionally independent, given a subset of nodes (excluding the pair of variables being considered) in the graph. A random variable X is conditionally independent of Z given another variable Y if $P(X|Y) = P(X|Y, Z)$. Intuitively, this means that Z does not

contain information about X that was not already contained in Y. Thus Z cannot be a direct cause of X. Let us assume that this conditional independence holds for X, Y and Z in Figure 2.4. If the variables are conditionally independent, then a necessary condition for causality is violated. Thus, the edge between the two variables can be eliminated (Figure 2.4, blocks 3 and 4 where the edge between X and Z gets eliminated). *A great majority of potential connections will be eliminated this way, leaving only a small set of potential causal relationships*.

The final step of learning the DAG representation involves determining the directionality of the edges such that the edges are directed from cause to effect (Figure 2.4, block 5). This can be done using orientation rules that are based on conditional independencies [20], and constraints that are based on the knowledge about the physical/temporal structure of the graph nodes. In the context of this research, we only use the temporal structure of the nodes to direct the edges, because we have found it to be the most reliable method for our applications [1].

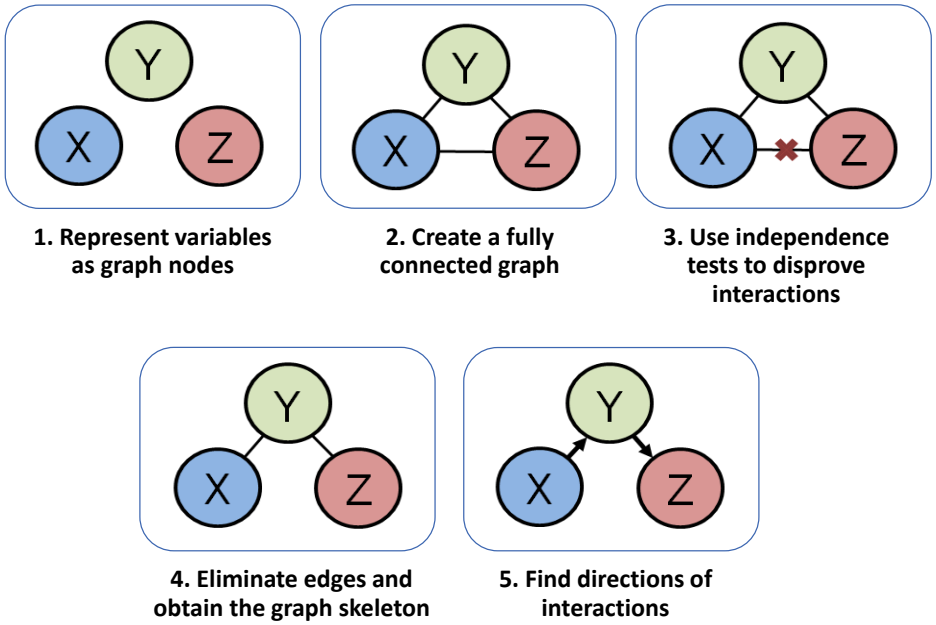


Figure 2.4: A step-by-step overview of the constraint-based structure learning methods.

PC Algorithm

The best-known algorithm for this purpose is the classic *PC* algorithm [37], which is named after its inventors, *Peter* Spirtes and *Clark* Glymour. As explained before, this method tests whether each pair of variables is conditionally independent, given any subset of nodes (excluding the pair of nodes itself) in the graph. The size of the subset used (i.e., the number of nodes in the subset) is referred to as the *order* of the conditional independence test. To make this process systematic and computationally efficient, the PC algorithm first checks whether each pair of variables connected in the graph are independent of each other (order = 0 test). If the variables are independent, then the edge between them will be removed. For the edges that remain after this step, the algorithm will test whether each pair of variables are conditionally independent, given another node in the graph (order = 1 test). The algorithm will then continue this process iteratively increasing the number of nodes included in the subset (increasing the order i.e., the cardinality of the conditioning set) until no more edges can be removed. This algorithm further improves the computational complexity of the conditional independence tests by only conditioning on the potential parents of the two variables. In our research, we use an order-independent variation of the PC algorithm, namely the PC-stable [38], which is more robust and easier to parallelize.

Temporal Extension of the PC Algorithm

Furthermore, as the causal interactions in climate span over time, we need to use a temporal model that is capable of capturing causal interactions that span over time. The PC, PC-stable and many other graphical causal model search approaches were originally introduced for cases that did not involve time series variables. Here we use a temporal extension of the *PC stable* algorithm following [1, 40], by additionally incorporating lagged copies of the original time series variables as nodes of the graphical model.

Figure 2.5 shows an example of a graphical model based on the temporal variables X, Y and Z. We provide the temporal information of the variables to the algorithm to ensure that the past and the present can cause the future, but not the other way around. The variables at the same time step

construct a *tier* of variables. In Figure 2.5, graph nodes belonging to the same tier are enclosed within a dashed brown color box. This tier information is used to determine the directionality of the graph edges. (We do not allow edges within the same tier, i.e., we do not allow instantaneous cause-effect relationships.) As discussed in detail in [1], the first few tiers in the temporal model will have a limited past to condition on, possibly resulting in initialization issues. A solution to this issue is to discard the first tiers and their connections.

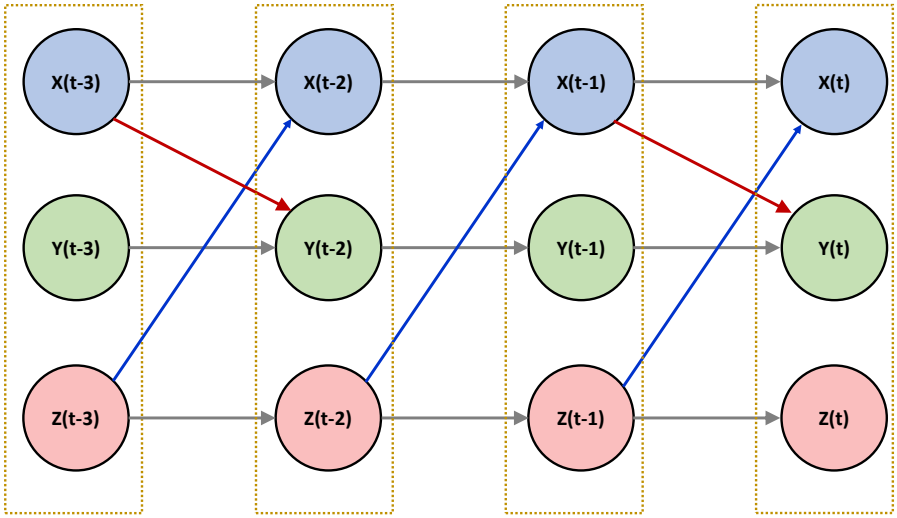


Figure 2.5: An example of a temporal model containing lagged variables as nodes of the graphical model.

As the final step, we identify interactions that are consistently repeating over the temporal causal structure. In Figure 2.5, there is a connection $Z \rightarrow X$ that occurs at a time delay of 1 time step. In the temporal model shown in the figure, this interaction repeats 3 times as $Z(t-3) \rightarrow X(t-2), Z(t-2) \rightarrow X(t-1)$ and $Z(t-1) \rightarrow X(t)$. This is the maximum number of times this interaction could have occurred in this specific model. However, we see that another connection that is $X \rightarrow Y$ at a time delay of 1 time step only occurs twice in the model. If there is a robust causal signal between two variables, we would expect that interaction to be consistent and repeat many times in our temporal model. However, if an edge does not show a repetitive nature and pops

up arbitrarily, it may indicate a false discovery. We present the final results as a summarized causal model as in Figure 2.6 where the directed edges between the graph nodes indicate the presence of potential causal interactions. The curved arrows indicate auto-correlations. The numbers alongside the arrows indicate the delay of the interactions in the number of time steps.

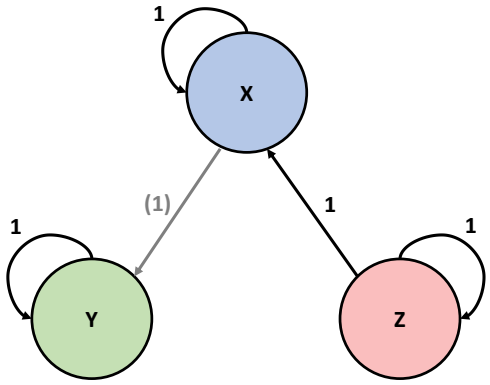


Figure 2.6: Summarized causal graph of the DAG presented in Figure 2.5. The grey edge represents that the connection from $X \rightarrow Y$ is less consistent.

Implementations

The method explained above is a simple temporal extension of the original PC-stable algorithm and is explained in detail in [1]. The author of this dissertation developed a Python based step-by-step tutorial suitable for low-dimensional applications that is provided in the Data Analysis Tools for the Atmospheric Sciences (DATAS) gateway [24]. The 'TETRAD' program [60], which provides a suite of methods and conditional independence tests is another powerful tool that can be useful to arrive at graphical causal models. In addition, the 'TIGRAMITE' [61] package provides tools to derive graphical causal models using linear as well as non-parametric conditional independence tests suitable for discrete or continuously-valued time series. The 'Bayes Net Toolbox (BNT)' [62] in MATLAB and the 'pclag' [63] package in *R* provide implementations of the PC stable algorithm for low dimensional applications. An efficient implementation in *C* developed by our group is used for the high dimensional application in Chapter 5.

Independence test based on partial correlation:

Partial correlation measures the linear association between two random variables, with the effects of a set of controlling random variables removed. For multivariate Gaussian variables that are linearly related, a zero partial correlation is equivalent to conditional independence [20]. We can test whether the partial correlation, ρ, is equal to zero by statistically testing $H_0 : \rho = 0$, using an appropriate statistical test such as the Fisher's Z-test [20]. Unless otherwise noted, we use this approach to determine conditional independence with PC-stable.

3 A Study of Links Between the Arctic and the Midlatitude Jet-stream[1]

This case study investigates causal links between Arctic temperatures and the midlatitude circulations. With respect to the circulations, we focus specifically on jet-streams, which are narrow bands of fast-moving winds. We apply two different frameworks for this application based on the concepts of (1) Granger causality and (2) Pearl causality. We use VAR-Granger and LASSO Granger methods under the Granger causality framework and the PC-stable algorithm under Pearl causality. This case study showcases how these methods can be useful to determine the time spans and strengths of the potential causal interactions between variables. It also portrays the similarities, pros and cons of these methods for low-dimensional applications. The motivation of this case study is to increase awareness of causal inference approaches that are easy to use and thereby facilitate the reader in questions related to method selection in low dimensional settings. The results from these methods show that Arctic temperature and jet speed and position all exhibit strong autocorrelation, but also that these variables are linked together by two robust positive feedback loops that operate on timescales of 5-25 days.

3.1 Motivation

Arctic amplification–that is, the phenomenon of Arctic temperatures rising much faster than the global mean [64] – and its present and future effects on midlatitude weather and climate have received substantial attention in recent years. While it is well known that the midlatitude circulation can drive changes in Arctic temperatures and sea ice, it is unclear how and to what extent the Arctic influences midlatitude weather [65]. Some argue that Arctic amplification is already

[1]This chapter contains material that has been published in the *Environmetrics* as: S. M. Samarasinghe, M. C. McGraw, E. A. Barnes, and I. Ebert-Uphoff: A study of links between the Arctic and the midlatitude jet-stream using Granger and Pearl causality. *Environmetrics*, **30:e2540**, `https://doi.org/10.1002/env.2540`.

influencing midlatitude weather [66–69], while others state that any possible signal is too small to have been observed amidst the background of atmospheric variability [70–72]. Regarding Arctic influence on the midlatitude circulation under climate change, idealized and fully-coupled climate model simulations have shown an equatorward shift of the jet streams and storm tracks (the regions where extratropical cyclones occur most frequently) as well as weakening of the zonal winds in response to Arctic warming and sea ice loss [73–76]. The mean locations of the wintertime jet streams in MERRA-2 reanalysis data [77] can be seen in Figure 3.1a. While this equatorward shift of the storm tracks is considered a robust midlatitude response to Arctic amplification, little is understood about the underlying dynamics behind this response in models or whether the models can adequately simulate the processes involved. Making progress requires that we study the two-way causal connections between Arctic temperatures and the midlatitude circulation–that is, the midlatitude circulation driving changes in the Arctic as well as the Arctic driving changes in the midlatitudes–in context of one another and the background of atmospheric variability.

The typical approach for assessing causal links in climate dynamics (including studying the links between the jet streams and Arctic warming/sea ice loss) is targeted modeling studies. While incredibly useful for understanding the physical mechanisms at play, this approach only allows for studying cause and effect in isolation, and does not allow for the feedbacks between the jet streams and Arctic warming to fully develop. A modeling study that examines the response of the atmospheric circulation to prescribed heating in the Arctic analyzes the circulation response to this Arctic warming; however, as the heating in the Arctic is fixed, the circulation is unable to influence, modify, or feed back upon the Arctic temperature. However, the midlatitude circulation is thought to have substantial impacts on the Arctic, potentially impacting Arctic temperatures, sea ice thickness and extent, and moisture [66, 78, 79]. These two-way feedbacks are potentially very important in terms of understanding the full atmospheric response to Arctic warming, and thus, it is vital to analyze both the Arctic's influence on the midlatitudes and vice-versa.

In addition, we have entered a period where atmospheric science tends to be "data rich" both in observations and model output [80]. Thus, there is great need for additional tools that can aid

scientists in identifying and extracting signals. Causal discovery techniques provide (1) robust definitions of causality, (2) can have direct ties to forecasting/prediction, (3) augment targeted model studies, and (4) allow for a direct comparison of results from observations and models.

Here we use two different frameworks to learn about causal relationships for the Arctic-midlatitude system. The first framework uses vector autoregression (VAR) type models (i.e. VAR and LASSO), combined with the concept of *Granger causality*. The second framework is based on the concept of *Pearl causality*. We apply both frameworks to the study of causal links between the Arctic and midlatitude jet streams. The purpose is two-fold: (1) by comparing the results of two very different frameworks we hope to identify robust results; (2) to make more geoscientists aware of the different types of causal analysis tools. In this article, we extend and progress from our initial study presented as a short paper [81], while providing detailed elaborations.

3.2 Causality in Climate Science

The best known concept to identify cause-effect relationships in climate science is *Granger causality*, developed by Clive Granger in the late 1960s [18,34] (see Section 3.4.1). However, other methods for causal discovery also exist, including *Pearl causality*, which was developed in the late 1980s [20,28,29,82] (see Section 3.4.2). When investigating cause-effect relationships it is crucial to distinguish between two different types of causal analysis, namely (1) intervention analysis and (2) observational analysis. *Causal calculus* [28, 82], which provides a comprehensive mathematical framework for causal discovery, analyzes what can or cannot be concluded using either type of analysis. Intervention analysis allows one to positively prove or disprove causal relationships between considered variables, but it requires that scientists can actually perform interventions on the system of study, i.e., that one can *change* the state of any variable and then observe the cascading effects on the other variables. In climate science, such interventions are very difficult and most often require the use of dynamical models and a specific experimental design [15].

Here, instead of performing multiple dynamical model simulations to act as interventions, we focus solely on an observational-type analysis, using model output that already exists in place of actual observations. Several different frameworks for observational analysis have been applied to cli-

mate science to provide graphical representations of likely cause-effect relationships [1,36,40–47]. However, of highest relevance to the application considered here are causality studies related to the Arctic, primarily the work by [83], [84] and [4]. These studies demonstrate the utility of causality techniques for studying Arctic-midlatitude connections. [83] and [84] use a vector autoregression (VAR) Granger approach to explore sea ice–North Atlantic Oscillation and sea ice–West Pacific Pattern connections, respectively. Their approach falls into the category of Granger approaches discussed in Section 3.4.1. While [83] and [84] assess the performance and significance of their VAR model and its results by comparing an unrestricted model (all connections are allowed) to a restricted model (specific connections are not permitted), here we assess significance by comparing an unrestricted VAR-Granger model (similar to that described in [83]) to a LASSO model that uses regularized regression to determine significant coefficients and connections (see Section 3.4.1). Another difference between our work and that of [83] and [84] is that these previous studies allow for simultaneous connections, while we - for simplicity - do not.

Kretschmer et al. [4] use a combination of Pearl causality and Granger causality concepts to investigate the role of Barents-Kara Sea ice concentrations on wintertime circulation using a 2-step procedure. In Step 1 they identify, for each considered variable, its potential causes using Pearl causality. Step 2 uses Granger causality, namely they perform a regression analysis for each variable to verify the results obtained in Step 1 and to assess the strength of the connections. Both steps are performed separately for each variable, i.e. one variable is declared the output variable (predictand) and the effect of all other variables (predictors) on the specific output variable is studied using scalar regression models. The results from these individual analyses are then pieced together at the end to a combined graph of all interactions. In contrast, we present using a vector approach that studies relationships between all variables simultaneously, without declaring any of them predictand or predictor, which allows for feedback loops between the variables to be explicitly identified. Nevertheless, the method by [4] is most closely related to the work presented here.

Another difference between our work and previous studies, e.g., [83,84] and [4], is that previous studies use empirical orthogonal functions (EOFs) to represent the circulation response to Arctic

warming. While EOF-based methods can be quite useful at dimension reduction and capturing broad-scale spatial variability, they often fail to capture details in the circulation changes. By focusing here on the position and strength of the jet streams separately, we are able to analyze potential circulation changes in more detail, as jet position and jet strength do not necessarily covary [85, 86].

3.3 Data

We analyze daily data from the Community Earth System Model–Large Ensemble (CESM-LE, [87]), a fully-coupled general circulation model (GCM). We use years 402 to 2,200 of the pre-industrial control run (all external forcing is fixed at its levels from 1850), resulting in 656,634 days (1,798 years) of data that acts as a proxy for a very long observational record. The data are gridded at a 1° horizontal grid spacing–that is, 0.9° in latitude by 1.25° in longitude. The seasonal cycle is removed from the daily data, and the daily data are then averaged into non-overlapping 5-day chunks to smooth out higher frequency variability. Afterwards, each time series is standardized–that is, its mean is subtracted and it is divided by its standard deviation. For our analysis we focus on the winter season (DJF), roughly dividing the number of data samples for each experiment by four.

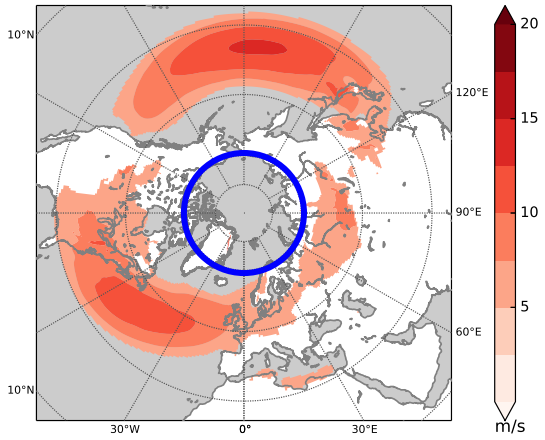
The data are stationary in the long-term, as they were generated by a control run with no external forcing; and the impacts of the seasonal cycle are minimized by removing the first four Fourier harmonics of the data. Nevertheless, as pointed out by [88], even if the seasonal cycle is removed in this way there is still the possibility that seasons affect the relationship between the different variables, which would necessitate taking the cycle into account when modeling their causal relationships. This is not a concern in this study, as we only consider a single season (Dec-Feb), and any seasonal changes can be assumed to be sufficiently small during that time. (Otherwise one could use the tools provided in [88] to test for periodic correlation, and to adjust the model accordingly). The mean subtraction in the standardization ensures that all variables used in the models have zero means. However, this step has minimal impact on the results as the variables already have means that are almost zero (but not exactly zero) at this stage. The winter

season is analyzed here because the teleconnection patterns between different regions are strongest in winter [89], and extratropical storm activity itself is greatest in the winter [90, 91]. Here, we focus on the North Pacific (120°E - 240°E, covering 97 grid boxes in longitude), and we analyze the circulation using the following three univariate time series, each consisting of 32,381 5-day averages:

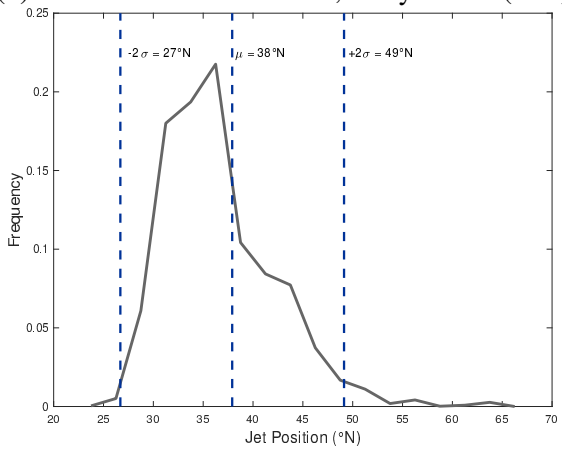
- jet latitude in the North Pacific, $\mathcal{L}$;
- jet speed in the North Pacific, $\mathcal{S}$;
- 850 hPa Arctic temperature averaged over 70°N-90°N at all longitudes, $\mathcal{T}$.

Jet latitude, $\mathcal{L}$, and jet speed, $\mathcal{S}$, represent the position and strength of the eddy-driven jet, which are often used as proxies for the position and strength of the jet-streams. $\mathcal{L}$ and $\mathcal{S}$ are calculated by determining the maximum position and strength of the zonal component of the winds at 850 hPa over the North Pacific basin (120-240°E). First, the 850 hPa zonal winds are averaged over all longitudes in the North Pacific basin to create a zonal-mean profile of the zonal wind. The resulting zonal-mean zonal wind profile is interpolated to a 0.01°latitude grid. A quadratic polynomial is fit around the maximum of this interpolated wind profile, following [92] (see Figure S3 in [93] for an example of a zonal-mean zonal wind profile). The maximum of this quadratic polynomial is the jet speed ($\mathcal{S}$), while its latitudinal position is the jet latitude ($\mathcal{L}$). The distributions of 5-day mean wintertime North Pacific jet position (Figure 3.1b) and jet strength (Figure 3.1c) can be seen in Figure 3.1, with their means and spreads indicated by vertical lines. Note that the seasonal cycle has not been subtracted from the jet position and jet speed distributions in Figures 3.1b,c, so as to show the actual physical values of North Pacific wintertime jet position and speed.

(a) DJF Zonal Winds, MERRA-2 (1980-2017)



(b) North Pacific Jet Position, 5 day means (DJF)



(c) North Pacific Jet Strength, 5 day means (DJF)

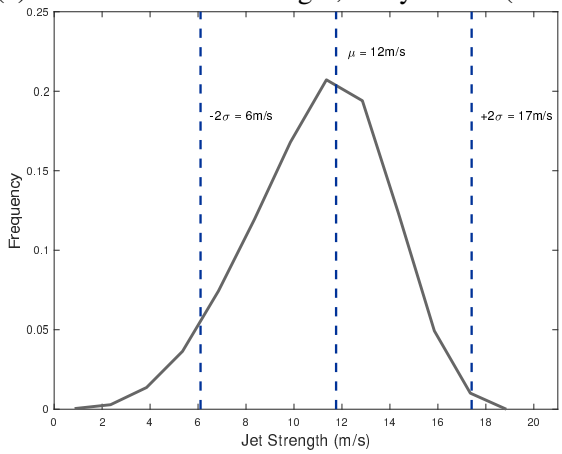


Figure 3.1: (a) Daily mean DJF zonal winds calculated from MERRA-2 reanalysis. Only values above 5 m/s are shown. The blue line denotes 70°N and indicates the cap over which $\mathcal{T}$ is calculated. Distributions of DJF North Pacific jet (b) position and (c) strength. Dashed lines indicate mean (μ) and two standard deviations ($\pm 2\sigma$).

3.4 Causality Methods

We briefly discuss the two types of causality methods used here, based on (1) Granger causality and (2) Pearl causality.

3.4.1 Methods Based on Granger Causality

In this subsection we discuss the concept of Granger causality, and introduce two types of regression models, *VAR* models and *LASSO* models, which can be used to identify variables that are *Granger causes* of other variables.

Granger causality

The concept of Granger causality is based on the predictability of temporal variables and can be simply explained using two univariate time series $X = \{x_t\}_{t=1}^{T}$ and $Y = \{y_t\}_{t=1}^{T}$. The time series X is said to be a *Granger cause* of Y, if the past values of X provide information about the current state of Y, beyond what is already known from the past values of Y, alone. Testing for Granger causality can thus be achieved by comparing two different prediction models. In the first model, Y is predicted using the past values of both X and Y (unrestricted model), whereas in the second model, Y is predicted using only the past values of Y (restricted model). Then a statistical test is carried out to determine whether the unrestricted model significantly improves the predictability of Y compared to the restricted model (for details, see for example [83]). If the predictability of Y is significantly improved by the inclusion of X in the model, X is said to Granger cause Y. As for the prediction model, a suitable linear or nonlinear model can be used. In the context of this paper, we only look at linear lagged regression models resulting in restricted and unrestricted models as follows,

$$\text{Unrestricted model:} \quad y_t = \beta + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_s y_{t-s} + c_1 x_{t-1} + c_2 x_{t-2} + \ldots + c_p x_{t-p} + e_u \tag{3.1}$$

$$\text{Restricted model:} \quad y_t = \gamma + b_1 y_{t-1} + b_2 y_{t-2} + \ldots + b_s y_{t-s} + e_r \tag{3.2}$$

Here a_i, b_i (for integer $i \in [1, s]$) and c_j (for integer $j \in [1, p]$) are scalar regression coefficients and β and γ are model intercepts. p and s are the number of lags of X and Y respectively. For simplicity, we use the same number of lags for each variable, i.e., $s = p$. e_r and e_u are the error terms of the restricted and unrestricted models respectively.

This approach, which we call the *two-model Granger* approach, is the simplest and by far the most common approach in which Granger causality is used in climate science to date. Use of Granger causality in this way constitutes a major step forward compared to more traditional methods in climate science such as single variable lagged linear regression analysis, since the latter can often lead to misleading conclusions, which are prevented with the two-model Granger approach [36]. The two-model Granger approach can be extended by connecting the concept of Granger causality to the coefficients of Vector Auto-Regression (VAR) models. This approach, which we call *VAR-Granger*, is a natural generalization of the two-model Granger approach, which allows us to (1) easily apply Granger analysis in cases with a large number of time series variables and (2) generate visualizations of the different relationships in the form of graphs, as demonstrated in later sections.

Vector Autoregression (VAR) model

Vector autoregression (VAR) performs regression on several variables *without* declaring one variable the predictand and the remaining ones predictors. Instead, all variables are treated the same way. Namely, a VAR(p) model estimates the vector $\mathbf{z}_t \in \mathbb{R}^k$, which contains *all* variables of interest, in terms of its p lags as follows:

$$\mathbf{z}_t = \mathbf{c} + \mathbf{A}_1\mathbf{z}_{t-1} + ... + \mathbf{A}_p\mathbf{z}_{t-p} + \mathbf{e}_t \quad \text{for} \quad t = (p+1), \ldots, T, \tag{3.3}$$

where vector $\mathbf{z}_t = [z_{1t}, \ldots, z_{kt}]'$ contains the values of k considered variables at time t;
$\mathbf{c} = [c_1, \ldots, c_k]'$ is a fixed vector containing the model intercepts; $\mathbf{A}_i$ are the $(k \times k)$ coefficient matrices (for $i = 1, \ldots, p$); and $\mathbf{e}_t = [e_{1t}, \ldots, e_{kt}]'$ is the vector of error terms (residuals). These error terms are assumed to be independent and identically distributed with $E[\mathbf{e}_t] = 0$ and a non-

singular covariance matrix $E[\mathbf{e}_t\mathbf{e}_t'] = \mathbf{\Sigma}_e$. Using this notation, to analyze the relationships between the two univariate time series X and Y from section 3.4.1, we would simply use $k = 2$ and define $\mathbf{z}_t = [x_t, y_t]'$.

An ordinary least-squares approach is used to solve the standard regression problem and calculate the model parameters of Eq. (3.3), namely vector $\mathbf{c}$ and matrices $\mathbf{A}_i$ [53]. Once a model of the form in Eq. (3.3) is obtained, we perform a validation test to ensure that the model is stable, i.e. that all roots of the characteristic polynomial lie outside the complex unit circle (e.g., see [53] or [94]). Such a stable VAR(p) process is automatically stationary as well. We derive such a VAR model for several different values of p, and select a suitable model using a selection criterion, such as the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) [53, 95, 96].

Connection to Granger causality

We can apply the concept of Granger causality to VAR models by inspecting the coefficients in $\mathbf{A}_i$. Let a^i_{lm} denote the element of row l and column m of matrix $\mathbf{A}_i$, and let the time-series $\{z_{l,t}\}$ denote the lth variable without lag, and $\{z_{m,t-i}\}$ denote the mth variable with lag i. Then a^i_{lm} denotes the effect of $\{z_{m,t-i}\}$ on $\{z_{l,t}\}$. Furthermore, since the data was normalized, a^i_{lm} indicates for a change of one standard deviation of $\{z_{m,t-i}\}$ the amount of change to expect in $\{z_{l,t}\}$. (This quantitative interpretation should be used with caution, as many geophysical relationships are nonlinear, and the model is thus only a rough approximation.) Then, for $l \neq m$, we see in this model that $\{z_{m,t-i}\}$ is useful for the prediction of $\{z_{l,t}\}$, if and only if $a^i_{lm} \neq 0$. Consequently, the mth variable, $\{z_m\}$, is said to *Granger-cause* the lth variable, $\{z_l\}$, if and only if at least one of the coefficients $a^i_{lm} \neq 0$ for any lag $i = 1, \ldots, p$.

For practical implementation, determining non-zero coefficients from a VAR model often requires applying a threshold on the magnitudes of the coefficients. That is because, due to noise and numerical accuracy, many - sometimes even most - of the coefficients can be very close to zero, but not exactly zero. Unless a natural threshold can be identified based on the problem, we need to set a user-defined threshold, resulting in subjective model selection, that can be highly sensitive

to noise. The following section describes how regularized regression techniques can be used to overcome this issue.

Regularized Regression (LASSO)

Regularization techniques can be used to obtain a sparse model that identifies the subset of predictors that have the strongest effect on the predictability of the responses. Specifically, the LASSO (Least Absolute Shrinkage and Selection Operator) [95, 97, 98] approach finds a least squares solution subject to a l_1-norm constraint on the coefficients, namely it imposes a bound on the sum of the absolute values of the coefficients. The imposed constraint shrinks the values of the regression coefficients and sets many coefficients exactly to zero to obtain a more generalized and sparse solution[2]. Therefore LASSO results in a model of the exact same form as Eq. (3.3), but where many coefficients are exactly zero. This improves the interpretability of the model (especially when there is a large number of predictors) and simplifies the process of identifying Granger causes. In addition, regularization allows for a more generalized model with improved prediction accuracy on test data compared to the ordinary least squares solution [97]. For this particular study, we use a version of the group LASSO approach, described in Appendix A.1.

For LASSO approaches, the regularization parameter, λ, determines the sparsity of the resulting matrix. Using $\lambda = 0$ is equivalent to the ordinary least squares solution, while a large value of λ gives a very sparse solution. (See Eq. (1) in Appendix A.1 for the exact definition of λ for Group Lasso.) λ should be selected as a trade-off, i.e. small enough such that all relevant relationships in the model are included, but large enough to yield a simple and interpreteable model. Literature suggests that the elastic-net, which is a regularization technique incorporating both l_1 and l_2 constraints, can be more useful than LASSO when predictors are highly correlated (which is often the case with many climate variables) [98]. However, in our study, the use of an elastic-net did not provide any noteworthy changes in the dependency structure compared to the LASSO results (therefore not presented here).

[2]Note that applying an l_2-norm constraint to the coefficients would also shrink them, but not force them to zero [98].

3.4.2 Method Based on Pearl causality

Pearl causality [20, 28, 29] follows from causal calculus [82] and is defined through interventions. If interventions are possible, then Pearl causality provides both *sufficient* and *necessary* conditions for causal relationships, i.e. with an intervention analysis one can say with *certainty* whether or not variable X is a cause of Y. In the case of observational analysis we can still use Pearl causality to provide a *necessary* condition for causality, namely certain conditions have to be satisfied in order for X to possibly be a cause of Y. The primary reason for losing the ability to prove sufficiency from observations alone is the potential existence of hidden common causes, i.e. *latent variables*. Namely, if two variables X and Y are both related solely due to the effect on them by an unmeasured third variable, Z, then an intervention analysis would correctly identify that there is no direct connection between X and Y, while an observational analysis might incorrectly conclude in that case that X is a cause of Y, or vice versa.

In an observational analysis, we can nevertheless use the necessary condition to eliminate the great majority of potential connections, leaving only a small number of potential cause-effect relationships. Namely, we use an elimination method that first assumes that all pairs of variables have cause-effect connections to each other (for all lags), then uses conditional independence tests as necessary causality conditions to *delete* connections. The best known algorithm for this purpose is the classic *PC* algorithm [37], which is named after the first names of the two authors, *Peter* Spirtes and *Clark* Glymour.

The PC algorithm usually yields a small set of *potential* cause-effect relationships, and the set of *true* causal relationships is a subset of that set. This type of method is called *constraint-based structure learning*, since we seek to constrain - i.e. construct a boundary set for - the set of true causal relationships. The specific method used here is the temporal version [1, 40] of the *PC stable* algorithm [38]. PC stable is a variant of PC that is more robust and can easily be parallelized. For more details on this approach, see [1]. For brevity, we refer to PC stable as simply PC in the remainder of this document.

3.5 Application to Arctic-Jet Stream Connections

To study the causal links between the Arctic and the midlatitude jet streams, we model the relationships between jet speed ($\mathcal{S}$) and Arctic temperature ($\mathcal{T}$), and, separately, the relationships between jet latitude ($\mathcal{L}$) and Arctic temperature ($\mathcal{T}$), using the three methods discussed above, i.e. VAR-Granger, LASSO-Granger and PC stable.

3.5.1 Implementation Details

Multivariate regression in VAR is implemented using the *mvregress* function in MATLAB, while the *Glmnet* [99] MATLAB package (multi-response Gaussian family scenario) is used to apply LASSO. We use the AIC and BIC criteria to select the maximum lag for the VAR model, p–here, we determine that $p = 5$ (25 days). The λ parameter in LASSO is selected using K-fold cross validation scheme (with $K = 10$) based on the mean square error performance on test data. Namely, the selected λ is the λ that corresponds to the *one-standard-error rule* as explained in [98, 100]–this λ gives a more regularized model, while keeping the mean squared error (MSE) within one standard error of the minimum MSE value. Based on this scheme, $\lambda = 0.0179$ is selected for the $\mathcal{T}$ vs. $\mathcal{S}$ model. (Additional information related to the selection of p and λ parameters is provided in Appendix A.2).

The *learn_struct_pdag_pc* function of the MATLAB Bayes Net Toolbox (BNT) [62] is used for a simple implementation of the PC stable algorithm. We modify the BNT to use temporal constraints [1, 40] and implement the PC stable version. We use the Fisher-Z test to test for conditional independence using partial correlation. The significance value, α, of the independence test is chosen as $\alpha = 0.05$, a common value; results are insensitive to variations in α between 0.01 and 0.1. Instantaneous connections are not permitted between variables in the PC stable algorithm to make results comparable with VAR and LASSO implementations. The PC model is run with 11 time slices, using the original variables $\{z\}$ and 10 time shifted versions of $\{z\}$ shifted by -25, -20, ..., -5, +5, ..., +25 days [1]. As general practice, we check convergence characteristics of the model to ensure that the model has converged to a solution [1].

3.5.2 Results and Interpretation

Jet Speed and Arctic Temperature

The VAR and LASSO models (Figure 3.2a,b) show both the magnitudes and the signs of the jet speed-Arctic temperature ($\mathcal{S}$-$\mathcal{T}$) relationship. For LASSO (Figure 3.2b), only the significant (nonzero) coefficients are shown, while all coefficients are shown for the VAR model (Figure 3.2a). Thus, lag days and coefficients that appear in both the LASSO and VAR models can be considered robust. To begin, we note that it is clear that the VAR and LASSO models are quite similar–the signs of the coefficients that appear in both models are largely the same, and their magnitudes are similar as well. The lags with the strongest coefficients are the same for both VAR and LASSO. For both the LASSO and the VAR models both $\mathcal{S}$ and $\mathcal{T}$ are autocorrelated (curved arrows), with coefficients that decay over the 25 day period but remain non-zero. $\mathcal{T}$ drives $\mathcal{S}$ 5 days later (as well as 10, 15, and 20 days later), with the positive coefficient indicating that warmer temperatures drive a faster jet in the North Pacific. $\mathcal{S}$ also drives $\mathcal{T}$ at a lag of 5 days, with the negative coefficient indicating that faster jets are associated with a colder Arctic. However, at a lag of 15 days and beyond, the relationship between $\mathcal{S}$ and $\mathcal{T}$ changes–$\mathcal{S}$ drives $\mathcal{T}$ with positive LASSO coefficients, indicating that a stronger North Pacific jet drives warmer Arctic temperatures. Collectively, the LASSO results indicate that there is a positive feedback loop between Arctic temperature and North Pacific jet speed–a warmer Arctic drives a stronger North Pacific jet, and the stronger jet drives further Arctic warming.

The jet speed–temperature PC model (Figure 3.2c) agrees quite well with the results of the LASSO and VAR models, although its formulation does not provide the magnitudes or signs of the relationships. The autocorrelation relationships (curved arrows) in the PC model are quite similar to those in the LASSO model. In the PC model, $\mathcal{T}$ drives $\mathcal{S}$ at a lag of 5 days only, and $\mathcal{S}$ drives $\mathcal{T}$ at lags of 15 and 20 days. These lags in the PC model match well with those having significant, larger coefficients in the LASSO models, highlighting that the different methods yield very similar (although not identical) results. The only exception is that for a lag of 5 days, LASSO picks up a moderately strong, negative connection from $\mathcal{S}$ to $\mathcal{T}$, while PC does not confirm that connection.

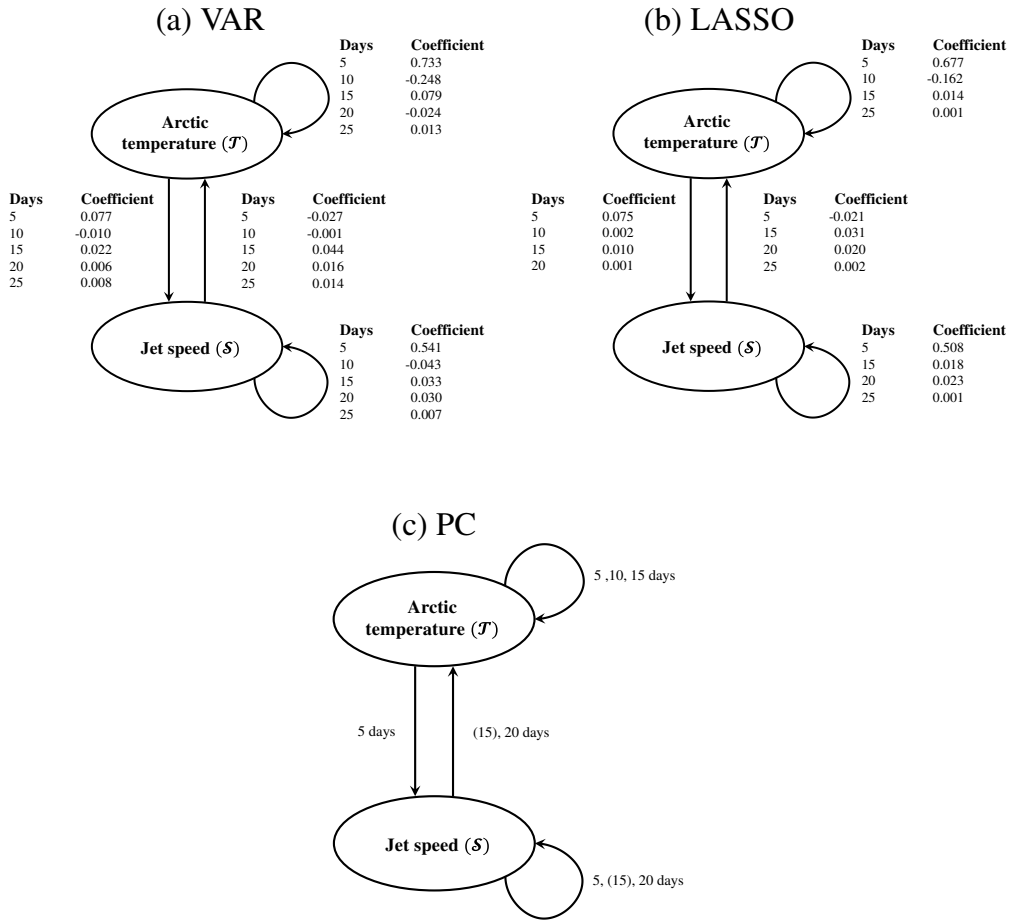


Figure 3.2: Arctic temperature ($\mathcal{T}$) and jet speed ($\mathcal{S}$) relationships as described by (a) VAR ($p = 5$), (b) LASSO ($\lambda = 0.0179, p = 5$) and (c) PC (11 time slices, $\alpha = 0.05$) models. Parentheses in PC results denote weak relationships.

Figure 3.4a summarizes these results–the colors of the arrows indicate the sign of the relationship (determined from VAR and LASSO), while the significant lag days are determined by including only lags that are significant in both the LASSO and PC models.

Jet Latitude and Arctic Temperature

Jet latitude, $\mathcal{L}$, also shows evidence of a causal relationship with $\mathcal{T}$ in the VAR, LASSO, and PC models (Figure 3.3). The influence of $\mathcal{T}$ on $\mathcal{L}$ is not as strong as the influence of $\mathcal{T}$ on $\mathcal{S}$, with both PC (Figure 3.3c) and LASSO (Figure 3.3b) showing few lags with significant relationships. In both the LASSO and VAR (Figure 3.3a, b) models, the $\mathcal{T}$ driving $\mathcal{L}$ relationship is negative at a lag of 5 days, and positive at a lag of 10 days. However, the PC model shows that the $\mathcal{T}$ driving $\mathcal{L}$

connection is only significant at 5 days; thus, we consider only the signs of the 5-day coefficients. This negative coefficient at a lag of 5 days indicates that warmer Arctic temperatures drive an equatorward shift of the jet in the North Pacific, consistent with results from previous modeling studies [73, 75, 76].

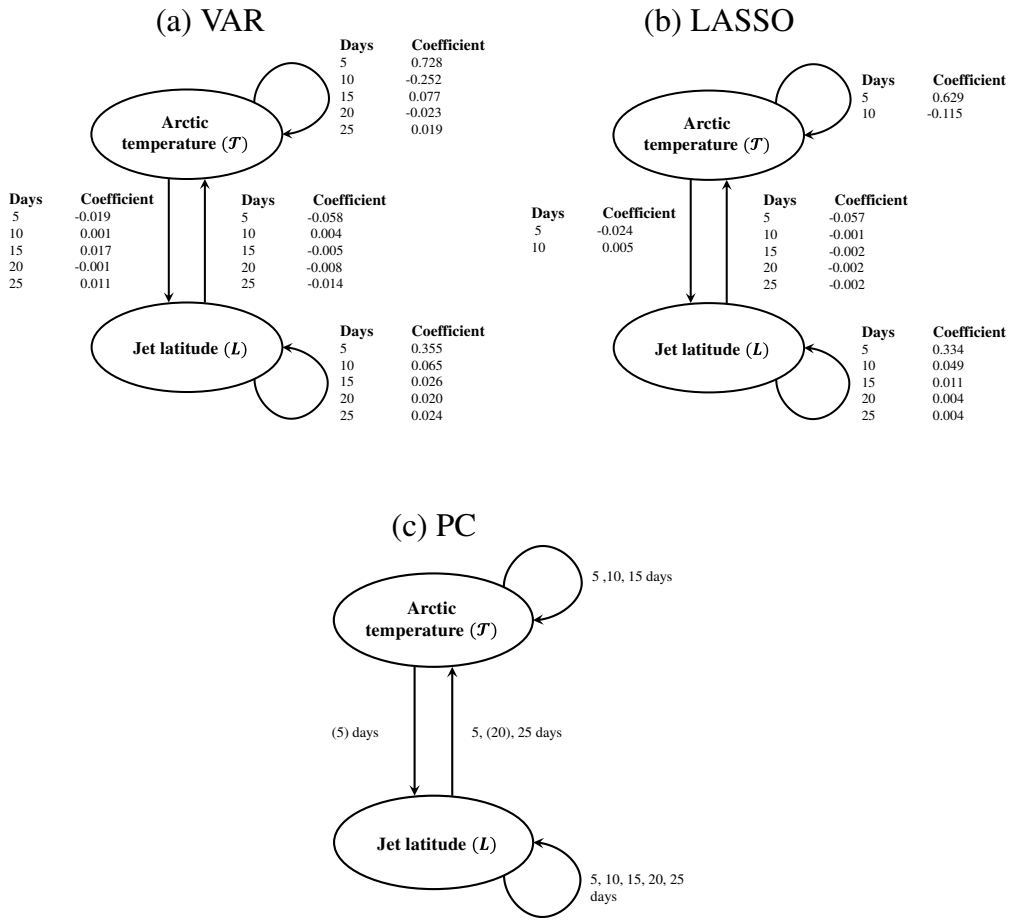


Figure 3.3: Arctic temperature ($\mathcal{T}$) and jet latitude ($\mathcal{L}$) relationships as described by (a) VAR ($p = 5$), (b) LASSO ($\lambda = 0.0337, p = 5$) and (c) PC (11 time slices, $\alpha = 0.05$) models. Parentheses in PC results denote weak relationships.

While the $\mathcal{T}$ driving $\mathcal{L}$ relationship is relatively weak, the influence of $\mathcal{L}$ on $\mathcal{T}$ is stronger. The LASSO and VAR models show that $\mathcal{L}$ drives $\mathcal{T}$ with negative coefficients at most lags, indicating that a more equatorward jet drives warmer Arctic temperatures and a more poleward jet drives cooler Arctic temperatures. This relationship is strongest at a 5 day lag; it also strengthens at a

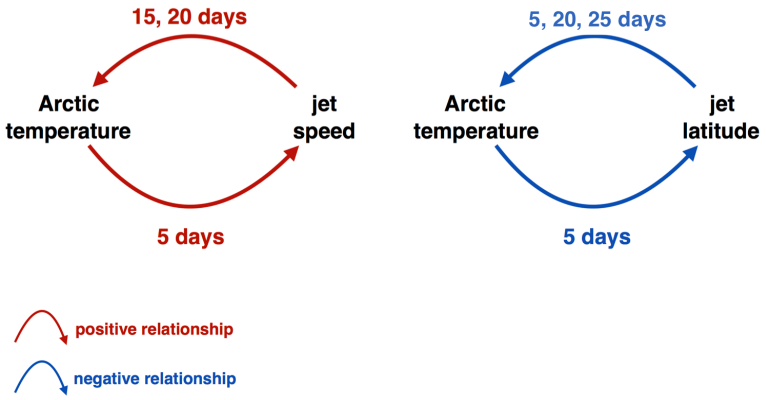


Figure 3.4: Minimalist graphs showing the dominant feedback loops and their time lags identified by all three methods when applicable. Note that the two positive relationships between arctic temperature and jet speed result in a positive (i.e. reinforcing) feedback loop. Similarly, the two negative relationships between arctic temperature and jet latitude together also result in a positive feedback loop.

25 day lag in the VAR model. The PC model shows a very similar relationship to the LASSO model–$\mathcal{L}$ drives $\mathcal{T}$ at a lag of 5 days, with a resurgence at 20-25 days.

Like the positive feedback loop found between $\mathcal{T}$ and $\mathcal{S}$ (Figure 3.2), the positive feedback loop in Figure 3.3 indicates that the jet position–Arctic temperature feedback loop also acts to enhance Arctic warming. This positive feedback loop between jet position and Arctic temperature is more simply represented in Figure 3.4b. Dynamically, a more equatorward jet could act to enhance Arctic warming via increased cyclonic Rossby wave breaking activity on the poleward flank of the jet stream [101], which increases moisture transport into the Arctic, and thus, atmospheric absorption of longwave radiation [78, 79], further warming the Arctic.

3.6 Discussion and Future Work

Using VAR, LASSO, and PC models, we have quantified robust positive feedback loops between the North Pacific jet stream and Arctic temperatures on daily timescales. Specifically, a positive feedback loop is identified between jet speed and Arctic temperatures on timescales of 5-15 days, where a warmer Arctic drives a faster jet, and a faster jet drives a warmer Arctic. Similarly, a positive feedback loop is identified between jet latitude and Arctic temperatures, where a warmer Arctic drives a southward shift of the jet which drives further Arctic warming.

The interesting similarity between results from different frameworks can be explained by the fact that partial correlations are used as the underlying indicator of relationships in all of these methods. In this study, PC uses partial correlation for its conditional independence tests while the Graphical Granger model uses regression coefficients, which are closely related to partial correlations as well. However, the results are not identical because the VAR-Granger models generate partial correlations by conditioning on all the variables simultaneously. PC performs partial correlation tests consecutively, starting from small conditioning sets and going to higher dimensions only as needed. As shown in [102], the PC approach has higher power compared to the VAR-Granger approach in high dimensional applications consisting of a larger number of variables and lags.

Throughout this study, we mainly focused on the linear relationships between Arctic temperature and the jet stream variables. We used linear regression models to identify Granger causes while we used partial correlation (a measure of linear dependence) based conditional independence tests in the PC algorithm to identify Pearl causes. However, the actual relationships in climate can be complex and nonlinear. By using nonlinear regression models (e.g., nonlinear random forests in [55]) along with statistical tests, one can identify nonlinear Granger dependencies between variables. It is possible that such nonlinear approaches would show higher sensitivity to Granger analysis, as they could pick up both the linear and nonlinear components of the causal relationship. Also, the use of entropy based measures (such as mutual information) for the conditional independence tests could capture nonlinear Pearl dependencies from the PC algorithm.

A limitation of any standard observational analysis, no matter whether based on Granger causality or Pearl causality, is the potential existence of hidden common causes. Results obtained from observation analysis must thus be interpreted with caution, i.e. they should be interpreted as *hypotheses* that need to be confirmed by identifying the physical mechanisms behind each connection. The Fast Causal Inference (FCI) algorithm [20], which is an extension of the PC algorithm, can be helpful for alleviating this issue, and is suggested to be studied in future work.

The work described here is only the beginning of a larger study that aims to apply causal discovery techniques to quantify Arctic-midlatitude connections. In future work we plan to expand the current study to include: (1) expansion of these methods to reanalysis, 2-D spatial fields, and inclusion of additional variables such as sea ice extent; (2) results from significance testing by comparing unrestricted and restricted VAR models; (3) quantification of the strength of causal relationships with non-regression-based techniques. By applying causality techniques to the topic of Arctic-midlatitude connections we aim to not only better understand the feedbacks between the Arctic and midlatitude weather, but also, provide the climate science community with an example of how these powerful tools can offer new insights into complex earth science problems.

Support for this work was provided by National Science Foundation (NSF) grants AGS-1545675 (Barnes and McGraw) and AGS-1445978 (Ebert-Uphoff and Samarasinghe) under the Climate and Large-scale Dynamics program.

4 Tropospheric and Stratospheric Causal Pathways between the MJO and NAO[4]

The Madden Julian Oscillation (MJO), which is a tropical oscillation that occurs above parts of the Indian and Pacific oceans, is well known as an influencer of the global atmospheric circulation and a primary source of predictability at subseasonal-to-seasonal (S2S) time scales. Previous work has shown that the Madden-Julian oscillation (MJO) can influence the North Atlantic Oscillation (NAO), which is an oscillation that influences North Atlantic (i.e., Europe and parts of North America) weather, via a Rossby wave teleconnection that propagates through the troposphere (i.e. a tropospheric pathway). In addition, recent work suggests that the MJO can influence the stratospheric polar vortex which is also known to influence the tropospheric NAO – thus, there likely exists a stratospheric pathway for MJO influence as well. Here, we apply two methods to shed more light on the pathways linking the MJO to the NAO. First, we use a traditional approach in climate science based on analyzing conditional probabilities. Second, we use methods from causal discovery theory based on probabilistic graphical models. Together, these two analysis approaches reveal that the MJO can impact the NAO via both a tropospheric and stratospheric pathway. The study and the results are explained in detail in the Sections 4.1 – 4.6 while Section 4.7 draws relationships between the two analysis approaches (i.e., conditional probabilities and graphical causal models). Finally, Section 4.8 provides a discussion of the practical challenges and considerations related to this case study.

[4]This chapter contains material that has been published in the *Journal of Geophysical Research: Atmospheres* as: E.A. Barnes, S.M. Samarasinghe, I. Ebert-Uphoff, and J. Furtado: Tropospheric and Stratospheric Causal Pathways between the MJO and NAO. *Journal of Geophysical Research: Atmospheres*, **124**, 9356-9371, `https://doi.org/10.1029/2019JD031024`.

4.1 Introduction & Motivation

Wintertime weather regimes across the Northern Hemisphere are generally typified by the phase and amplitude of tropospheric modes of variability, like the Northern Annular Mode (NAM) (e.g., [103, 104]) or its more regional expression, the North Atlantic Oscillation (NAO) (e.g., [105–107]). In the positive phase of the NAM, the polar jet stream, and subsequently the storm track, is located anomalously poleward of its mean position, and equatorward during the negative phase. Changes in the jet stream position also govern wintertime regimes, with cold conditions prevailing in the Northern Hemisphere mid-latitudes during negative NAM regimes (e.g., [104, 108]). As such, understanding drivers and dynamics of subseasonal-to-seasonal (S2S) variability in the jet stream is an active goal for wintertime S2S forecasting applications (e.g., [109–111]).

In the last two decades, two large-scale modes of climate variability have been studied extensively as precursors to changes in the Northern Hemisphere midlatitude jet streams. The first is stratosphere-troposphere dynamical coupling in the Northern Hemisphere extratropics, particularly that associated with variability in the stratospheric polar vortex, an area of low pressure that resides 10-50 km above Earth's surface in the Arctic during the boreal cold season. Theory, observations, and model experiments demonstrate that stratospheric circulation anomalies descend with time and can affect the tropospheric circulation with same-signed anomalies (e.g., [112–116]). For example, for weak stratospheric polar vortex conditions (i.e., when climatological westerlies weaken or potentially reverse to easterly), anomalous easterlies descend from the stratosphere into the troposphere, weakening the zonal component of the tropospheric polar jet. Consequently, tropospheric eddies reorganize and positively feed-back onto these tropospheric jet anomalies (e.g., [116–119]), shifting the polar jet stream equatorward and inducing a negative NAM regime for the next 2-8 weeks [115]. Therefore, changes in the dominant circulation pattern within the Northern Hemisphere wintertime stratosphere (i.e., the stratospheric polar vortex) project onto changes in the dominant pattern of the extratropical tropospheric circulation (i.e., the NAM). As such, the stratosphere may indeed play a role in long-lead predictability of the Northern Hemisphere midlatitude circulation [111, 120, 121].

In addition to stratospheric variability, the Madden-Julian Oscillation (MJO) [122] has also been investigated for its links to variability in the hemispheric NAM, as well as a similar regional mode of North Atlantic jet stream variability, the North Atlantic Oscillation [105–107]. Mechanistically, tropical convection associated with the MJO excites Rossby wavetrains in the upper troposphere that propagate from the tropics into the midlatitudes of both hemispheres (e.g., [123–125]). For the Northern Hemisphere in particular, these Rossby wavetrains interact with the polar jet stream, modifying its strength and position across the hemisphere (e.g., [126–128]) and impact associated teleconnection patterns. As such, the interactions between the MJO and the NAM and NAO are currently largely considered to act via a tropospheric pathway. Indeed, studies have illustrated changes in the phase and amplitude of the NAO associated with different phases of the MJO. Several studies (e.g., [129–132]) show that the positive phase of the NAM and NAO typically follows strong (i.e., high-amplitude) MJO Phases 2, 3, and 4, with negative phases of the NAM and NAO following MJO Phases 7 and 8. Using a statistical model, [133] shows significant 2-week lead predictions of both phases of the NAO when using the phase and amplitude of the MJO as a predictor. Similar predictability was found when looking at output from the Global Environmental multiscale (GEM) model for connecting the MJO to the NAO [134]. Other studies suggest that the connections between the MJO and the NAO may also be a function of other modes which affect the tropospheric waveguide, including the phase of the El Niño-Southern Oscillation (ENSO) (e.g., [135, 136]) and/or the Quasi-Biennial Oscillation (QBO) (e.g., [137]).

Moreover, the MJO is also known to have impacts on the state of the Northern Hemisphere stratospheric polar vortex. These impacts can occur either via vertically-propagating Rossby waves directly associated with the MJO convection or indirectly from MJO-extratropical tropospheric teleconnections that produce preferential wave patterns that excite planetary-scale waves which subsequently propagate into the polar stratosphere (e.g., [138–142]). Indeed, the MJO has been used as a skillful predictor for Northern Hemisphere sudden stratospheric warming events [143, 144]. These changes in the stratospheric polar vortex can then dynamically alter the tropospheric circulation and the NAO, as discussed above. Therefore, complexity exists in the

interactions between the NAO, the MJO, and the Northern Hemisphere stratospheric polar vortex - as each is thought to have distinct connections with the others. So, while previous work has suggested that the MJO's impact on the NAO is via a tropospheric pathway, it also seems possible that the MJO's influence on the NAO could also arise because of its interactions with the stratospheric polar vortex (i.e. a stratospheric pathway). Alternatively, the stratospheric polar vortex may instead condition the background flow of the Northern Hemisphere extratropics to be more or less conducive for MJO-related wave propagation, thus also impacting the MJO-NAO connections. Furthermore, in all of these cases, there remains the question of timing. For example, how many days does it take for the signal from the MJO to reach the NAO? These questions are difficult to answer via traditional statistical analyses (e.g., lagged regressions and composites) given the continuous interplay among the three phenomena, as well as the confounding issue of memory, or autocorrelation, in the system.

Here, we apply two methods to shed more light on the relationships between the MJO, the stratospheric polar vortex, and the NAO. First we use a traditional approach in climate science based on analyzing conditional probabilities. Second, we use methods from causal discovery theory based on probabilistic graphical models to identify tropospheric and stratospheric pathways linking the MJO to the NAO, distinguishing different pathways following different phases of the MJO. This method allows us to also quantify the time lags associated with these different causal connections. Supported by the causal discovery results, we then interpret the conditional probabilities to determine whether the stratosphere acts as a mediator, or serves as the pathway itself, in modulating MJO-NAO interactions.

4.2 Data & Approach

4.2.1 Data

We analyze three different daily, wintertime indices in this work, one for the Madden-Julian oscillation (MJO), one to capture variations in the strength of the stratospheric polar vortex (VORTEX) and one to represent the state of the North Atlantic tropospheric circulation (NAO). For

all three indices, the time period spans from 1979-2016. We define the extended winter season as November-March (NDJFM) and center all of our analysis on these months. However, since lagged-relationships are at the center of this study, lagged variables are allowed to extend into October and April for the causal discovery method (Section 4.4). For the conditional probability analysis (Section 4.3), the MJO index is restricted to October-February which allows the VORTEX and NAO indices to extend into March and April when appropriate.

The MJO index is that of the Australian Bureau of Meteorology (BoM), and is based on the real-time multivariate index (RMM) of [145]. This index captures the location and strength of anomalous tropical convection as well as the larger-scale winds. The RMM MJO index is composed of two separate indices, RMM1 and RMM2, which represent the two leading principal components, obtained using extended empirical orthogonal function (EEOF) analysis of the combined fields of tropical (15°S-15°N) outgoing longwave radiation and the 200 hPa and 850 hPa zonal winds. As is the convention, the amplitude of the RMM MJO index is defined by $\sqrt{RMM1^2 + RMM2^2}$ and the angle between RMM1 and RMM2 defines the phase (values 1-8).

The VORTEX index is computed from daily stratospheric geopotential heights from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis [146]. All analysis is performed for the VORTEX defined at 100 hPa, and then again defined at 50 hPa. The 100 hPa is of particular interest here as we are focused on the coupling component between the stratosphere and the troposphere (e.g., [147]), and thus, this level is more likely to couple with the troposphere. The 50 hPa definition constitutes a more classical definition of the Northern Hemisphere polar vortex / stratospheric circulation variability used in several other works (e.g., [104, 114, 148, 149]), and thus allows us to bolster our results with this definition. The geopotential heights are area-weighted averaged over the polar cap, defined as 65^oN-90^oN, and daily anomalies are computed by subtracting the daily climatological value over the 1979-2016 period. The values are then standardized by dividing by the standard deviation of all daily values between November-March, and finally multiplied by -1 to stick with our convention of a positive index (i.e. VORTEX >1) denoting a stronger polar vortex.

The daily NAO index is that of the NOAA Climate Prediction Center [150], and is based on the Rotated Principal Component Analysis of [151] applied to 500 hPa geopotential height anomalies. Anomalies are based on the 1950-2000 climatological daily mean and standard deviation. Two missing values were found in the original NAO index downloaded from CPC, and we replace these two missing values with the values from the day before. This has no effect on our results or conclusions.

Both the VORTEX and NAO indices are smoothed by applying a backward 5-day running mean to the time series in order to capture variability beyond synoptic timescales, with the last day of the 5-day average used to denote the date. We do not smooth the MJO data. To increase our sample size within each MJO phase we combine adjacent phases together (e.g. 2/3, 4/5) as is often done in the literature (e.g., [129, 152]). Sample sizes for Figures 3-6 are provided in the Appendix B.2.

For the conditional probability results in Section 4.3, the three indices are further simplified by turning them into discretized values. Specifically, MJO active periods are defined as days when the MJO amplitude is greater than 1.0, and days when the MJO amplitude is less than 1.0 are considered to have no MJO activity. The VORTEX and NAO indices are labeled "1" if their indices are above 1.0, labeled "0" if between -1.0 and 1.0, and labeled "-1" if below -1.0, unless otherwise specified. For the causal discovery results in Section 4.4, the MJO index is also split into active and inactive periods, but the NAO and VORTEX indices are left continuous.

4.2.2 Causal Discovery Method

Causal discovery methods based on probabilistic graphical models [20, 28] have recently been used in climate science to identify potential cause-effect relationships [1, 3, 40]. They yield graphical representations, such as the one in Fig. 4.1, where each variable is represented as a node (circle) and arrows between the nodes represent identified potential cause-effect relationships. There are two key facts to know about such methods. (1) The methods seek to identify only *direct* causal relationships. For example, consider the situation indicated in Fig. 4.1, where variable X is a cause of Y, indicated by the arrow $X \to Y$, and Y is a cause of Z, indicated by the arrow $Y \to Z$. There

is no arrow directly from X to Z, indicating that the only causal connection between X and Z is through Y. The connection from X to Z (through Y) is called an *indirect* causal connection, and an important feature of causal discovery is to distinguish between direct connections (shown as arrows) and indirect connections (no arrow shown). (2) Such methods can only identify *potential* cause-effect relationships, that is, the relationships can be true causal relationships, due to latent (i.e. hidden) variables, or a combination of both. Thus the purpose of these methods is to generate *hypotheses* of causal relationships for further testing, *not* to prove such relationships.

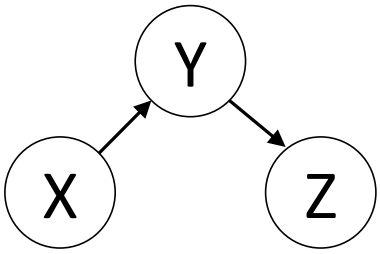


Figure 4.1: Causal graph representation of three variables, X, Y, and Z, that form a causal chain from X to Y to Z. X is a *direct* cause of Y, Y is a *direct* cause of Z, but X is only an *indirect* cause of Z (through Y).

The type of method used here is called constraint-based structure learning and is an elimination process that yields as output a graph structure of the type shown in Fig. 4.1, where the nodes represent the variables of interest and the arrows, or edges, indicate potential direct causal connections. The method starts with a full graph, where all nodes are assumed to be connected to all other nodes. Edges in the graph are then pruned by using conditional independence tests. If two nodes in the graph are conditionally independent provided a set of other nodes in the graph, then these two nodes cannot have a direct interaction between them. In such a case, the initially assumed connection between the two nodes is eliminated. The specific algorithm used to obtain the graph structure is called *PC stable* [38], a modification of the classic *PC* algorithm [20]. Furthermore, we use a temporal extension, where in addition to the regular variables we introduce lagged copies of these variables as separate nodes to allow modeling causal relationships that take one or more days from cause to effect [40, 153]. Once the edges are established through conditional independence

tests, the edge directions are established through temporal constraints in the temporal model (i.e. a cause can only happen before its effect). For details on the use of the temporal extension, see [1]. Section 4.8 summarizes some of the assumptions and limitations of this method.

A complete probabilistic graphical model consists of a graph, plus probability tables, namely one for each node, to form a complete statistical model - thus the name “probabilistic" graphical model. The interested reader can find a discussion of these probability tables, and of their relationship to the probability tables developed in Section 4.3, in Section 4.7. Note that within the scope of this paper we are only interested in the graph part of the probabilistic graphical models, for reasons discussed in Section 4.7.

In this approach we use the three variables, VORTEX, NAO and MJO, and, employing the temporal extension discussed above, additionally include several lagged copies of each variable, which in this case are chosen to be 5 days apart. Because we expect that the state of the MJO might change the causal relationships, we create four different causal graphs, namely one for each MJO phase pair, i.e. phases 2/3, 4/5, 6/7, 8/1. Developing separate models for each phase pair provides the freedom to identify significant differences in the causal structure that may exist between different phases. If no significant differences exist, then these four models should be very similar. Each of the four models represent VORTEX and NAO as continuous variables, and MJO as a binary variable. The MJO variable is set to 1 if the MJO index has an amplitude greater than 1.0 and is in the relevant phase range for that model, and 0 otherwise.

To determine conditional independencies, we statistically test for zero partial correlations using Fisher’s Z test. This test assumes linearity of relationships and Gaussianity of data as discussed in Sections 4.8. In this study, we use this common conditional independence test as a first approximation of the causal relationships although some of the variables are binary. We do this because existing independence tests for mixed variables have the limitation that their performance has not been well tested for real world applications [154], making us less confident in their outputs. Nevertheless, we did try the Conditional Correlation Independence (CCI) and the Conditional Gaussian Likelihood Ratio tests for this study. However, we do not present results here as the outputs were

sensitive to the specific parameter choices used by the different methods. The results using Fisher-Z are presented in Section 4.4. Specific parameter choices are described in Appendix C.

4.3 Results: Conditional Probability Tables

Before diving into the more complex method of graphical models, we begin with a standard approach in climate science, namely, conditional probabilities. Specifically, we compute conditional probabilities of the state of the NAO conditioned on the state of the VORTEX and the MJO indices at various prior lags (e.g., the probability that the NAO index is 1 conditioned on the MJO being in phase 2/3, fifteen days before). We introduce the following notation: NAO(t) denotes the NAO value at time t. Similarly, VORTEX$(t-5)$ denotes the VORTEX state 5 days before the NAO at time t. Since conditional probabilities by themselves are not based on causality theory - in contrast to our constraint based learning results in Section 4.4 - one must be careful when interpreting high or low probabilities as causal. With that said, these conditional probability calculations provide strong evidence of the MJO driving the NAO weeks later, and allow us to decipher whether the stratosphere acts as a mediator of this MJO-NAO link and/or a pathway itself.

To demonstrate our methods before launching into the full results, Figure 4.2 shows the probability that the NAO$(t) = 1$ given that the MJO was active 15 days prior (MJO$(t-15) > 1$) for varying VORTEX states 5 days prior (VORTEX$(t-5)$). This choice of lagging the VORTEX by 5 days is motivated by results from the causal discovery method presented in Section 4.4. The y-axis denotes the probability that NAO$(t) = 1$, while the x-axis denotes the MJO phase 15 days prior, and the two panels denote VORTEX defined at different pressure levels. Different colors denote the different values of VORTEX$(t-5)$. The dashed lines denote the probability that NAO$(t) = 1$ given the state of the VORTEX$(t-5)$ when the MJO is inactive (i.e. the RMM amplitude is below 1.0; or $P(NAO(t) = 1|VORTEX(t-5), MJO(t-15) < 1)$). Hence, the values are constant across the different phases of the MJO. These lines act as a baseline about which to assess the relationship between the MJO and the NAO 15 days later. For example, regardless of the MJO, the NAO has a higher probability of being positive when the VORTEX is positive 5 days before (dashed orange line) compared to when the VORTEX is negative (dashed purple line) or neutral

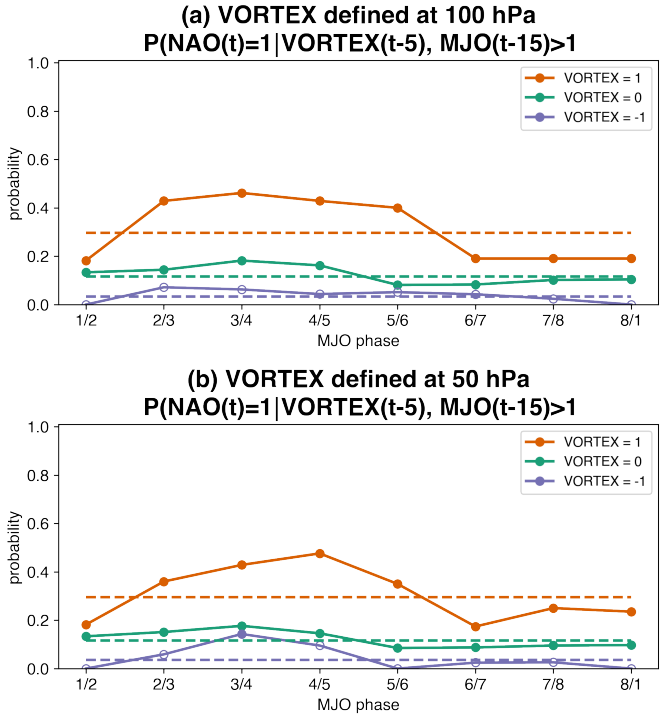


Figure 4.2: The probability (or fraction of days) that the NAO$=1$, fifteen days following different MJO phases and given the state of VORTEX five days prior. Different colors denote different VORTEX states, and dashed lines denote the same NAO probability except for non-active MJO periods, i.e. $P(NAO(t) = 1|VORTEX(t-5), MJO(t-15) < 1)$. Open circles denote combinations that occur only 1 or 2 days in the entire record.

(dashed green line). This is consistent with established literature on stratosphere-troposphere dynamical coupling in the Northern Hemisphere (e.g., [114, 115, 155–157]).

Focusing next on the subdivision by MJO phase and the case of VORTEX$(t-5) = 1$ (solid orange line), we find that the NAO has a much higher probability of being positive 15 days following MJO phases 2/3 to 5/6. Increased probability of the NAO$(t) = 1$ (in comparison to an inactive MJO; dashed lines) is also seen following these general MJO phases when VORTEX$(t-5)$ is neutral (green line), but is weak and infrequent when VORTEX$(t-5)$ is negative (purple line; open circles denote combinations that occur only 1 or 2 days in the entire record). This will be further discussed below. Similarly, the NAO has a much lower probability of being positive 15 days following MJO phases 6/7 to 8/1 and 1/2.

The important comparison to make in Fig. 4.2 is the *difference between the solid and dashed curves*, which quantifies the *additional* information about the future NAO state provided by the

MJO. That is, we know from previous work that the stratospheric polar vortex can drive changes in the NAO (e.g., [157] and references therein), and thus, it is the additional information provided by the MJO beyond this baseline (dashed lines) which we wish to quantify. In what follows, we show the same types of results, only we plot the difference between the solid and dashed curves (MJO active vs MJO inactive) as a function of time lag.

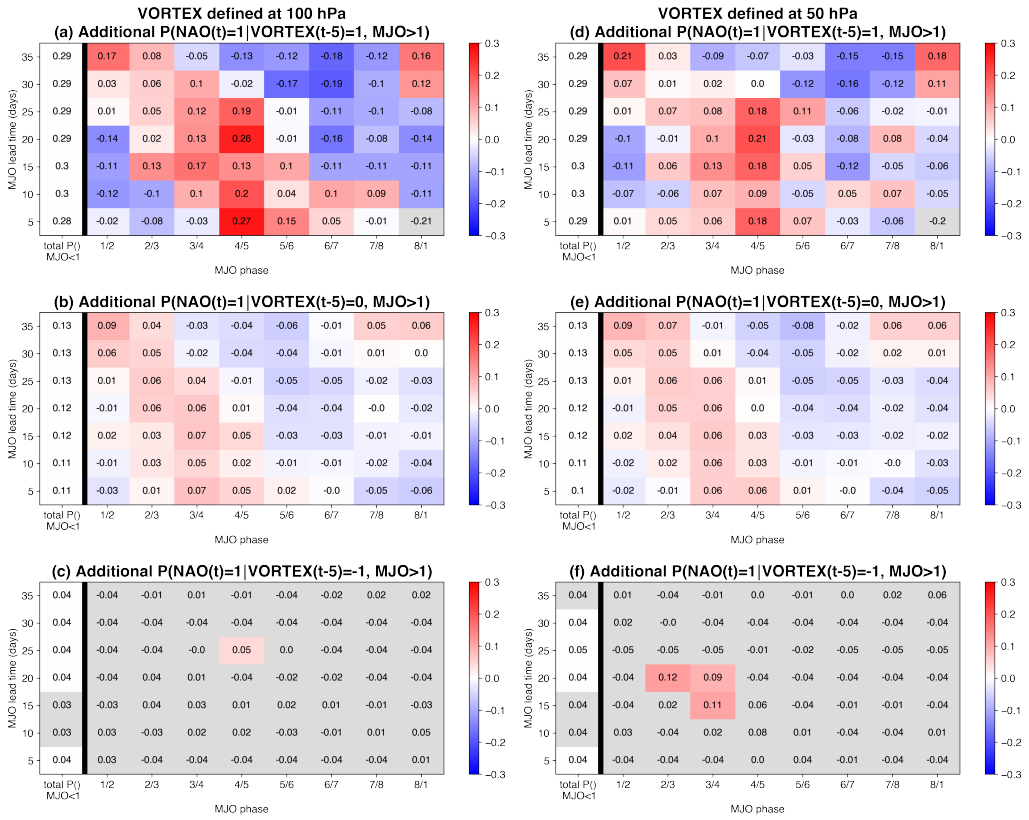


Figure 4.3: The additional probability (beyond that when the MJO is inactive; as shown in the left-most column) that the NAO $= 1$ for various MJO lead times and phases given that (a) VORTEX $= 1$, (b) VORTEX $= 0$, and (c) VORTEX $= -1$. All VORTEX states are for 5 days prior to NAO $= 1$. That is, the $P(NAO(t) = 1|VORTEX(t-5), MJO(t-T) > 1)$ where T is the label on the y-axis. Printed numbers in each box denote the value, and combinations that occur fewer than three days are grayed out. Note that unlike the other columns, the left-most column value denotes the *total* probability when the MJO is inactive, i.e. $P(NAO(t)|VORTEX(t-5), MJO(t-T) < 1)$.

Figure 4.3 shows the the additional (i.e., beyond that when the MJO is inactive) probability that NAO$(t) = 1$ as a function of MJO lead in days (y-axis), and MJO phase along the x-axis, for different states of VORTEX 5 days before. As a reminder, the orange line in Fig. 4.2 can be thought

of as a slice at an MJO lead time of 15 days along Fig. 4.3a, except what is shaded in Fig. 4.3 is the difference between the solid and dashed lines. A quick glance at Fig. 4.3 reveals something remarkable, namely, the propagation of the MJO can be seen in the probabilities as stripes in MJO phase/lead space for both VORTEX definitions. Stripes appear from the upper-left corner of the panels and propagate down and to the right. This striping shows a very clear modulation of the NAO probability by the MJO and supports the notion that the conclusions we make are built on a physical relationship between the NAO and MJO.

Fig. 4.3 shows results of the probability that the NAO$(t) = 1$ given that (a,d) VORTEX$(t-5) = 1$, (b,e) VORTEX$(t-5) = 0$ and (c,f) VORTEX$(t-5) = -1$. From these panels, we see that the MJO is only able to cause changes to NAO(t) when the VORTEX state aligns. That is, when VORTEX$(t-5) = 1$ (Fig. 4.3a,d), the MJO increases the probability that NAO$(t) = 1$ by 20-30%, the greatest seen when for MJO Phases 4/5 and leads up to 20-25 days. This is almost a doubling from the total probability when the MJO is inactive, which is approximately 30% as seen in the left-most column (Fig. 4.3a,d). When VORTEX$(t-5) = 0$ (Fig. 4.3b,e), NAO probability changes are also nonzero, demonstrating that the NAO responds to the MJO without the VORTEX; however, the additional probability values are not as large as for VORTEX$(t-5) = 1$. Therefore, the VORTEX being in a positive state may precondition the troposphere to better respond to MJO teleconnections. Finally, when VORTEX$(t-5) = -1$ (Fig. 4.3c,f), the odds of the NAO$(t) = 1$ are weakly altered (especially when VORTEX is defined at 100 hPa) with only a few instances when such conditions occur (gray shading in Fig. 4.3c,f denotes combinations that occur fewer than three times throughout the record), further demonstrating that the VORTEX has a strong influence on whether the NAO is able to respond to the MJO. When VORTEX is defined at 50 hPa there is increased probability of the NAO being positive when VORTEX is negative (4.3f), however, only 3 days go into each of these averages (see Appendix B.2 Figure B1).

The picture thus far is that, while the MJO can influence the NAO at up to 30+ day leads, it can only do so when the stratospheric vortex is in a state that is conducive to the NAO responding. That is, it can only do so when the future VORTEX and resulting NAO indices do not have

opposing signs. To further demonstrate this, Fig. 4.4 shows similar plots to Fig. 4.3, but now for the probability of the NAO(t)$= -1$. In this instance, the same relationships are found but for the opposite signs. A propagating signal is found when VORTEX$(t-5) = -1$ (Fig. 4.4c,f) and when VORTEX$(t-5) = 0$ (Fig. 4.4b,e). However, when VORTEX$(t-5) = 1$ (i.e., a strong stratospheric polar vortex), in direct opposition to an NAO$(t) = -1$ (Fig. 4.4a,d), there is no evidence of an influence of the MJO on the NAO as these conditions rarely occur.

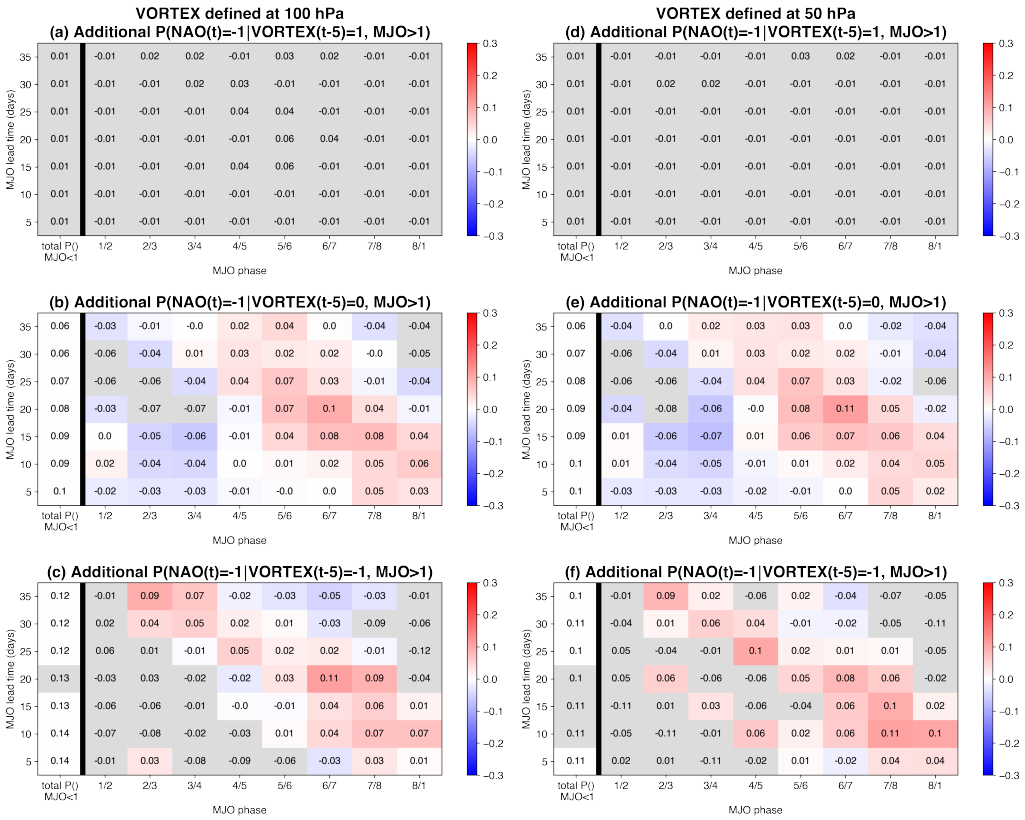


Figure 4.4: As in Fig. 4.3 except for NAO$(t) = -1$.

The magnitudes in Fig. 4.4c,f (VORTEX$(t-5) = 1$)) are smaller than in Fig. 4.3a,d (VORTEX$(t-5 = 1)$). This may be due to the different response of the troposphere to different flavors of weak vortex events (e.g. displacements vs splits) [158] whereas most strong vortex events are similar to one another. This would likely cause the composite anomaly magnitudes under VORTEX$(t-5) = 1$ to be larger due to this larger consistency in the tropospheric response.

Note that our analysis thus far does *not* preclude that the NAO itself is causing the VORTEX to be of a certain sign, rather than the response of the NAO to the MJO somehow being modulated. This can occur since we are not considering the autocorrelation of the NAO here - e.g., NAO$(t-10)$ could affect both VORTEX$(t-5)$ and NAO(t) (e.g., [36]). However, as we will show in Section 4.4, the causal discovery analysis when VORTEX is defined at 100 hPa suggests that there is a direct causal connection between VORTEX $\rightarrow$ NAO but *no direct connection* of NAO $\rightarrow$ VORTEX, implying that our worry of NAO $\rightarrow$ VORTEX is likely not valid there. When VORTEX is defined at 50 hPa, a direct connection between NAO $\rightarrow$ VORTEX, however, is identified for some MJO phases.

Before moving to the causal discovery method, we present another set of results where we treat the NAO as a continuous variable (i.e. we do not discretize it to be -1, 0 or 1). In this instance, we construct plots of the average NAO index (rather than conditional probabilities) following certain MJO phases. This compositing allows us to visualize changes in the continuous NAO index regardless of whether the NAO changes discrete states. Figure 4.5 shows the results where now the shading denotes the average NAO index at time t. Once again, we see stripes which represent the propagation of the MJO and give us confidence that the NAO behaviour is indeed physically linked to the MJO. When VORTEX$(t-5)=0$ (Fig. 4.5b,e), we see a stripe of positive NAO indices following phases 2,3,4 and negative NAO indices following phases 6,7,8, consistent with the findings of, e.g., [129] and [133]. When VORTEX$(t-5)=1$ (Fig. 4.5a,d), the shading is almost all positive, which shows once again that the NAO(t) tends to be positive when the VORTEX$(t-5)$ is positive. However, the NAO is even more positive in a stripe starting soon after phase 4/5, suggesting that certain phases of the MJO are able to make the NAO more positive, or, using our previous results, more positive more often. We do not find a strong negative NAO stripe here since the VORTEX$(t-5)=1$ condition makes a negative NAO regime highly improbable, fitting nicely with our earlier results. When VORTEX$(t-5)=-1$ (Fig. 4.5c,f), the shading is almost all negative, indicating that the NAO tends to be negative when the VORTEX$(t-5)$ is too. However, like with the positive case, we see a very negative NAO stripe starting soon after phase 8, and in

quadrature with the stripe seen in panel (Fig. 4.5a,d). Taken together, Figs. 4.4 and 4.5 present two complementary views for looking at how the MJO may drive changes in the NAO based on the state of the stratospheric vortex. Namely, the NAO appears insensitive to the MJO when the VORTEX anomaly is of the opposite sign to that of the expected NAO response.

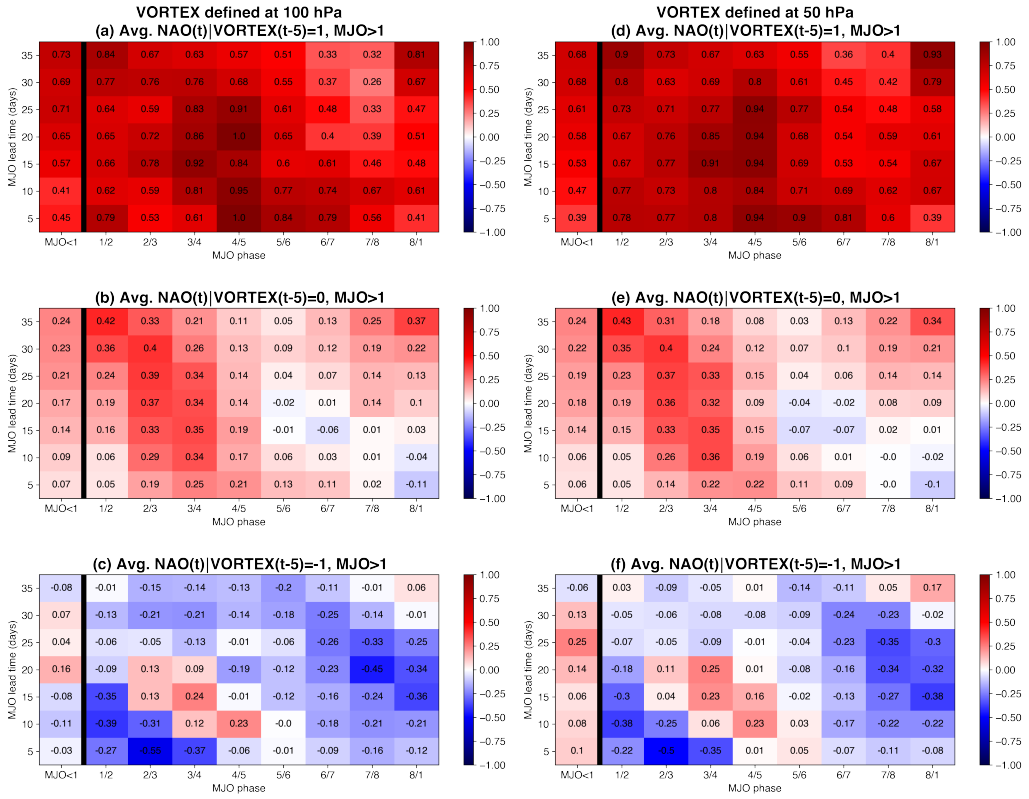


Figure 4.5: The average NAO index for various MJO lead times and phases given that the VORTEX is in a certain state 5 days prior. That is, the average $NAO(t)|VORTEX(t-5), MJO(t-T) > 1$, where T is the label on the y-axis. The left-most column shows the average NAO index when the MJO is inactive. Printed numbers in each box denote the value, and combinations that occur fewer than three days are grayed out.

One can follow a similar method to investigate the influence of the MJO on the VORTEX index itself. Such an analysis is shown in Figure 4.6, where VORTEX is now allowed to be continuous (as we did for the NAO in Fig. 4.5) and we plot the average VORTEX index for different MJO phases and lead times. The results agree well with previous works (e.g., [139, 141, 144]), namely that negative VORTEX indices (weakened vortex state which can lead to a sudden stratospheric

warming) tend to follow MJO phases 6/7, while positive VORTEX indices tend to follow MJO phases 2/3.

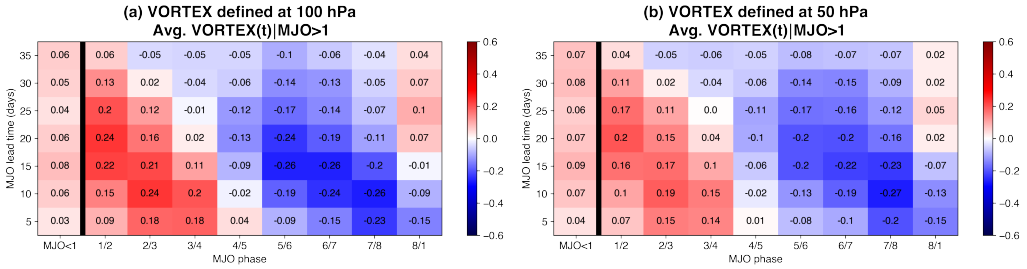


Figure 4.6: As in Fig. 4.5 except for the average VORTEX at various lags following different phases of the MJO.

As stratospheric polar vortex fluctuations are known to drive tropospheric NAO responses (see Section 4.1), the question arises whether the influence of the MJO on the future state of the NAO (as presented in Figs. 4.3-4.5) is due to the MJO's influence on the vortex which *then* influences the NAO (i.e. a stratospheric pathway), the MJO's *direct* influence on the NAO (i.e. a tropospheric pathway), or a combination of the two. Furthermore, once a causal pathway is established, what is the time delay between cause and effect? To address these questions we turn to causal discovery methods and present the direct connections in the form of graphical models in the next section.

4.4 Results: Causal Graphs

To augment our conditional probability approach in Section 4.3, we additionally provide results using a much less common approach, namely, the causal discovery method described in Section 4.2.2. Figs. 4.7 and 4.8 show the four graphs obtained for MJO phases 2/3, 4/5, 6/7 and 8/1. These graphs summarize the potential causal relationships identified with the temporal extension of the PC stable algorithm. The advantage of these summarized graphs is that, for each MJO phase pair, they provide a big picture view of all potential *direct* causal connections between the three variables, including corresponding lag times, auto-correlation and potential feedback loops. This summary of possible causal connections cannot be achieved using the conditional probability

approach presented in Section 4.3, primarily because these conditional probability tables cannot distinguish between direct and indirect connections.

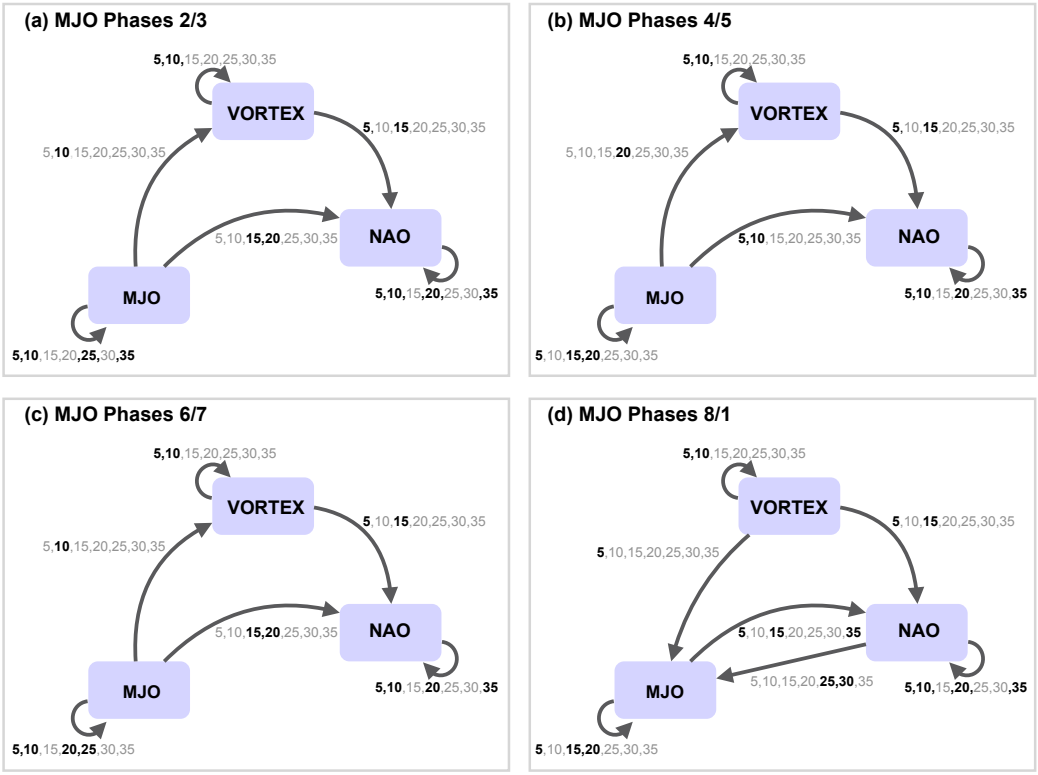


Figure 4.7: Summary graphs of the causal links identified for MJO phases (a) 2/3, (b) 4/5, (c) 6/7, (d) 8/1 for VORTEX defined at 100 hPa. Black numbers (in bold) denote the lag (in days) of the identified connections, while no significant causal connections are detected at any of the other lags (shown in gray). Arrows that loop back on themselves denote auto-correlation. Thin lines denote connections present using a 90% confidence threshold, while thick lines denote connections that are also present using a 95% confidence threshold. All connections here are significant at 95% confidence.

To illustrate the meaning of these graphs, consider Fig. 4.7a, which indicates causal relationships when the MJO is in phase 2 or 3. This graph indicates that VORTEX can causally affect NAO primarily at a time delay of roughly 5 and 15 days (black numbers). That is, a change of the VORTEX state may cause a change of the NAO 5 days later. Likewise, MJO Phases 2/3 may causally affect the NAO at a delay of 15-20 days, with no strong connections at any of the other lead times (Fig. 4.7a, gray numbers). Auto-correlations are indicated by loops that point from each variable back to itself and are generally consistent across MJO phases.

Continuing our focus when VORTEX is defined at 100 hPa (Fig. 4.7), the graphs for MJO Phases 2/3 (Fig. 4.7a) and 6/7 (Fig. 4.7c) are almost identical, the only difference being in the significant time lags of the auto-correlation of the MJO. This result is somewhat expected from the MJO literature, in that phases 2/3 and 6/7 produce similar, but opposite-signed, teleconnection responses (e.g., [129, 159, 160]). The graph for MJO phases 4/5 (Fig. 4.7b) is quite similar in overall structure to phases 2/3 and 6/7, although the time delays for several of the connections differ. That is, in all three graphs shown in Fig. 4.7(a)-(c), causal connections between variable pairs go from MJO to VORTEX, from VORTEX to NAO and from MJO to NAO. In contrast, the MJO phases 8/1 graph includes connections from NAO to MJO and from VORTEX to MJO that are not identified for any of the other phases. Finally, the stratospheric pathway (MJO to VORTEX) is not present in the 8/1 graph, however, the graphs for phases 1 and 8 individually suggest that this stratospheric pathway is present for phase 1 alone (not shown). The smaller sample size for single phases, however, makes us less confident in the resulting graph.

Causal graphs were also produced for VORTEX defined at 50 hPa, and results are shown in Fig. 4.8. Conclusions are very similar to those at 100 hPa, except that additional causal links are retained for the VORTEX at 50 hPa case. Namely, for phases 2/3, 4/5 and 6/7, a causal link exists between the NAO to the VORTEX at lag of 5 days. Phases 2/3 also exhibit a link between the NAO to MJO. Other than these four extra connections and slight differences in the time delays, the causal graphs for VORTEX defined at 50 hPa and 100 hPa are identical. With that said, some of the connections are slightly weaker at 50 hPa compared to 100 hPa (thin lines in Fig. 4.8) since they no longer appear when a significance threshold of 95% (instead of 90%) confidence is used. Causal graphs for VORTEX defined at 70 hPa were also computed (not shown), and the connections are nearly identical to those at 100 hPa except for extra connections from the NAO to VORTEX for all MJO phases. All connections for 70 hPa are significant at 95%, suggesting that the slightly weaker connections found at 50 hPa may be due to slightly weaker coupling between 50 hPa and the troposphere.

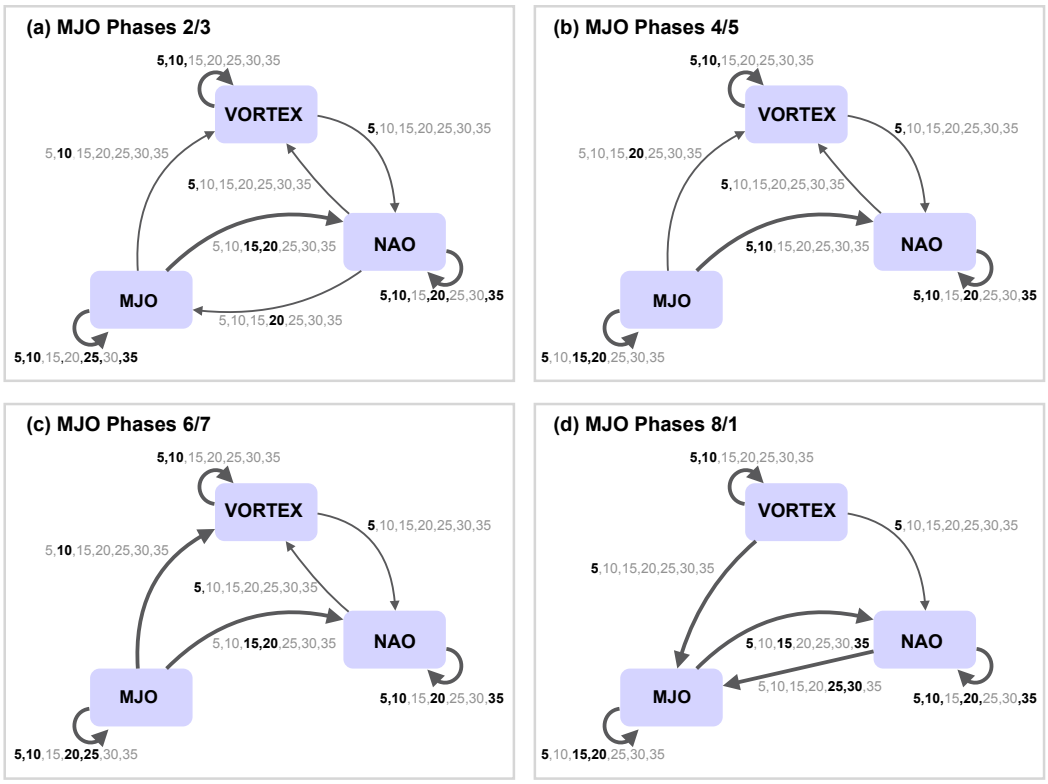


Figure 4.8: Summary graphs of the causal links identified for MJO phases (a) 2/3, (b) 4/5, (c) 6/7, (d) 8/1 for VORTEX defined at 50 hPa. Black numbers (in bold) denote the lag (in days) of the identified connections, while no significant causal connections are detected at any of the other lags (shown in gray). Arrows that loop back on themselves denote auto-correlation. Thin lines denote connections present using a 90% confidence threshold, while thick lines denote connections that are also present using a 95% confidence threshold.

A causal connection between the NAO to the MJO (Fig. 4.7d and 4.8d) is actually supported by the literature. [130] show that the NAO can influence the tropical winds in the Atlantic-African sector at lags of 10-25 days through anomalous wave propagation, and this time delay is consistent with what we find here (Fig. 4.7d). They argue that the anomalous winds in this region could act to trigger or amplify the MJO. Our results here suggest that the direct causal influence of the NAO is on MJO phases 8/1 at a lead of 25-30 days.

Modulation of MJO phases 8/1 by the stratospheric polar vortex is less supported by previous work (Fig. 4.7d and 4.8d). However, several studies have examined how weak vortex episodes (e.g., sudden stratospheric warmings) impact tropical tropospheric and lower stratospheric temperatures. [161] and [162] show that, for sudden stratospheric warmings events, the mean merid-

ional circulation features enhanced upwelling across the tropical tropopause, effectively cooling the lower stratosphere in the tropics. This enhanced circulation changes the vertical temperature structure which can alter the static stability of the upper troposphere/lower stratosphere, which impacts tropical convection anomalies. However, as MJO activity is also modulated by different phases of ENSO (e.g., [136, 163, 164]) and different phases of the QBO (e.g., [137, 165]), disentangling connections from high-latitude processes to tropical convection remains difficult.

We emphasize that the graphs in Figs. 4.7 and 4.8 only include arrows for *direct* connections, but indirect connections can easily be inferred. For example, in Fig. 4.7a, we see that a change in the MJO might have a strong effect on the NAO 15-20 days later, but due to auto-correlation of both the MJO and NAO, there may be indirect effects that last much longer. In fact, according to Fig. 4.7a, the MJO may be highly predictive of the NAO state not only after 15-20 days, but even 5, 10, 20 and 35 days *after* those initial 15-20 days, due to autocorrelation of the NAO (not to mention the additional autocorrelation of the MJO). Thus, the purpose of these graph representations is not to derive predictive models since this would require one to consider the combined effects of direct and indirect relationships. Instead, their purpose is to shed light on the true causal pathways that are responsible for the combined effects, and for that, one has to untangle the direct and indirect connections as done here. In contrast, the detailed conditional probability tables in Section 4.3 are ideal for the purpose of prediction, as they incorporate the direct and indirect connections into one probability.

4.5 Discussion

The conditional probability tables in Section 4.3 model the combined effect of direct and indirect connections from the MJO and the stratospheric polar vortex to the NAO and reveal how combinations of specific vortex states and MJO phases at certain lead times relate to the different states, and amplitudes, of the NAO. In contrast, the causal graphs in Section 4.4 distinguish between direct and indirect connections, but at the price of losing information about the signs of the relationships and the propagation of the MJO across its different phases. Thus, the graphical models extract the direct causal connections across the considered variables, while the conditional

probability tables allow us to explore the nature of the relationships. By combining results from these two analyses we are able to draw conclusions about the existence and nature of a tropospheric and/or stratospheric pathway linking the MJO to the NAO.

Specifically, we find that the there are likely direct connections from the MJO to the stratospheric polar vortex (strongest at a time delay of 10 or 20 days), the stratospheric polar vortex to the NAO (strongest at a time delay of 5 days) and the MJO to the NAO (with varying time delays - depending on MJO phase - between 5 and 30 days). There is also a robust link from the vortex to the MJO in MJO Phases 8/1. Thus, there are two clear causal pathways from the MJO to the NAO: (1) A direct one, and (2) an indirect one from the MJO through the stratosphere and then downward to the tropospheric NAO. While each individual connection has been explored individually in the literature, our analysis shows that the full pathways are supported by formal methods from causality theory and places them in the larger context of all possible causal connections across the three variables. In addition, this work has identified that the stratospheric polar vortex can further influence the NAO by acting as a "gate" that does or does not let the causal effect of the MJO pass to the NAO.

4.6 Conclusions

We investigate the causal pathways via which the MJO and the stratospheric polar vortex can influence the NAO on subseasonal-to-seasonal timescales using two distinct, but related, methods. The first method follows a standard approach in climate science, whereby one creates conditional probability tables based on the different combinations of the MJO, stratospheric polar vortex, and the NAO. While this method helps identify the signs and strengths of the connections, it inherently combines the effects of direct and indirect connections, including the influence of autocorrelation of the different phenomena. The second method, based off of causal discovery theory, identifies the direct connections via probabilistic graphical models and is able to identify specific time lags of these direct connections but does not quantify the sign of the connection.

While each of these two methods has its limitations, together, they make a powerful pair. Both methods show that the MJO impacts the NAO via a direct tropospheric pathway, but also confirm

a distinct, indirect, stratospheric pathway. Specifically, the MJO influences the strength of the stratospheric polar vortex on a timescale of $\sim$10 days, and then 5 days later the vortex drives changes in the NAO. The conditional probability tables also demonstrate that the stratospheric polar vortex conditions the tropospheric circulation to be conducive (or not) to MJO influence. Specifically, the NAO is only able to respond to the MJO when the state of the stratospheric vortex supports the expected NAO response (i.e. the same polarity). When this is the case, the NAO is up to 30% more likely to be in a particular state following active MJO events. This knowledge could be useful in subseasonal forecasting frameworks for the NAO and thus improve our skill at predicting weather regimes across Europe and parts of North America for those timescales.

These results rest on a large volume of work examining relationships between the MJO, the NAO and the stratospheric polar vortex, as discussed in the Introduction. Based on this previous knowledge, our work outlines how two different techniques can be combined to offer a more complete view of the causal pathways connecting these three climate phenomena. In fact, recent work has utilized a version of causal networks to predict the stratospheric polar vortex state [5], assess the influence of the stratosphere on cold extremes [166], and quantify specific aspects of stratosphere-troposphere coupling [4]. While this previous work, as well as our own, has applied causal discovery techniques to the stratospheric circulation, there is nothing particular about the stratospheric polar vortex, the MJO or the NAO that makes them particularly conducive to causal discovery methods. Thus, we fully expect these approaches to be highly relevant for a large array of climate dynamics research.

4.7 Relationships of Conditional Probability Tables to Graphical Models

As mentioned in Section 4.2.2, a complete probabilistic graphical model consists of a graph indicating the direct relationships, plus probability tables that indicate the probabilistic nature of these relationships. In Figures 4.7 and 4.8, we presented summarized versions of the graph part of the model. However, such graphs alone do not tell us anything about the nature of the connections, that is, the sign of the relationship, or whether the vortex acts as a mediator between the MJO and the NAO. To address these types of questions, the second step of learning a probabilistic graphical

model typically consists of calculating, for each node, the probability table given all of its potential causes (i.e. the *parents* in the graph) [20, 28]. These specific probability tables are useful as each node in a graphical model is conditionally independent of its nondescendants, given its parents. (For any node X_1 in the graph, another node X_2 is a nondescendant if there is no directed path from X_1 to X_2.) This property is also known as the *causal Markov condition* [20]. Utilizing this property, the joint distribution of n nodes, $X_1, \ldots, X_n$, can be expressed in a minimal fashion as (see [20, 58])

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\text{parents}(X_i)). \tag{4.1}$$

Eq. (4.1) decomposes the high-dimensional joint probability table (on the left) into a product of much lower dimensional conditional probability tables (on the right), by taking advantage of the underlying dependency structure encoded in the graph part of the graphical model. The key point here is that we know from the graph structure which variables can have a direct impact on which other variables, and only those need to be considered to express a node's conditional probability, $P(X_i|\text{parents}(X_i))$. The two parts of the graphical model, the graph and the conditional probability tables, $P(X_i|\text{parents}(X_i))$, assigned to each node X_i, thus provide a complete representation of the joint distribution.

Taking advantage of this minimal decomposition we can study the individual probability tables of each X_i to determine whether there are any interesting patterns/properties between the variables. For example, to study the nature of the impact of MJO and VORTEX on NAO, one would first study the probability table,

$$P(\text{ NAO(t) } | \text{ parents(NAO(t)) }), \tag{4.2}$$

where the parents of $NAO(t)$ denote the set of all nodes with an arrow pointing directly to NAO in the graphs (the set of potential direct causes).

Including connections from all four cases in Fig. 4.7 the set of all potential NAO causes is $S = \{$ VORTEX(t-5), MJO(t-5), MJO(t-10), MJO(t-15), MJO(t-20), MJO(t-35), NAO(t-5), NAO(t-10), NAO(t-20), NAO(t-35)$\}$. The probability table for NAO thus spans 12 discrete variables, i.e. it is still too high dimensional. Given limited sample size, studying the entries of such a table would not yield robust insights, as many table entries would rely on very few samples. In such a case researchers often marginalize out individual variables to reduce dimensionality further, i.e. conditioning only on subsets of S, to find interesting patterns. This is exactly what was done in Section 4.3 of this study. That is, we already know that some subsets yield interesting patterns, namely Fig. 4.3 and 4.4 show conditional probability tables for NAO for subset $S_{sub} = \{$VORTEX(t-5), MJO(t-T)$\}$ for different lags T. Thus, the conditional probability tables in Section 4.3 can be seen as a simplified stand-in for the high-dimensional conditional probability tables in Eq. (4.2) that illustrate important patterns relating VORTEX and MJO to NAO. (Note that even with this interpretation the simplified conditional probability tables in Fig. 4.3 and 4.4 do not distinguish between direct and indirect causes, because the auto-correlations - which are part of many of the indirect pathways in Fig. 4.7 - were marginalized out here.)

4.8 Additional Discussions with a Methods Perspective

As discussed in detail in the previous sections, this case study elucidates the causal pathways between the MJO and NAO by using graphical causal models alongside simplified conditional probability calculations. In a methods point of view, setting up the inference problem to answer the questions of interest to this research posed certain challenges. The first and biggest challenge in this case study involved setting up the inference problem in a physically meaningful way. As discussed in section 4.2.1, the MJO has a continuous amplitude, and a phase that is represented as a categorical variable with 8 cyclic states. With this, we came across the initial question of how to best represent the MJO in the learning algorithm. The dynamics of the MJO is known to change based on its phase, also the amplitude is indicative of the strength of the MJO's influence. Even though both the amplitude and phase carry information relevant to the MJO, creating a graphical model representing these variables as separate nodes can result in a complex model (for example,

with interactions such as the MJO phase causing the MJO amplitude or the MJO amplitude causing the NAO) that can be difficult to interpret. Therefore as discussed before, we introduce a binary indicator variable indicating an MJO event such that the information pertinent to the MJO can be represented using a single node in the graph. The indicator variable is defined such that MJO is 1 when there is an MJO event in the phase of interest, i.e., MJO amplitude is at least 1 and the phase is in the phase of interest to the study, and 0 otherwise.

With the binary encoding of the MJO variable, we now have mixed variable types - binary (MJO) and continuous (NAO and VORTEX), which brings out considerations related to suitable method selection. Creating graphical models for each of the 8 phases separately using this encoding also results in issues related to sample size. For each model, we only have a small number of samples relating to MJO events in the phase of interest (in other words, the number of samples with 1s is much smaller than the number of samples with 0s). As a correlation or partial correlation based conditional independence test only captures an average causal interaction between the variables, a small sample size relating to the events of interest can deter the method's ability to capture important relationships. To alleviate this issue, we decide to combine adjacent phases as 2/3, 4/5, 6/7 and 8/1 based on the recommendation of the domain expert. An alternative approach to the setup used in this research would have been to represent the MJO as a categorical variable with 5 states. Four states representing MJO events in phases 2/3, 4/5, 6/7, 8/1, and a fifth state representing MJO non-events. Even though this type of analysis would have been possible with the use of a mixed (continuous and categorical) graphical model, this would have required extensive analysis of conditional probability tables to extract all the important information encapsulated within the graph. As an extensive analysis of probability tables can be difficult, especially with a large number of parent nodes in a temporal model, we selected the simpler option to create causal models for each phase separately.

As mentioned before, the graphical models presented in the study only capture average causal interactions between the variables. Also, they do not unravel the specific states associated with the variables, e.g., whether the interaction is occurring at 1 or 0 of the MJO. Additionaly looking

into the conditional probability tables allow to clarify and get a deeper understanding of the mechanisms governing the relationships encapsulated through these compact graphical representations. Through this discussion we see that the seemingly simple process of setting up the causal inference problem actually requires careful consideration and thought. It also requires close collaboration with a climate scientist or other domain expert to ensure that the graphical models facilitate answering the science questions that we are actually interested in.

4.8.1 Assumptions and Limitations of the Graphical Model Approach

In this section we discuss a few limitations of the graphical model approach used here. It is common to use this approach in spite of these limitations, but one should be aware of these limitations for the interpretation of the results.

1) Linearity and Gaussian distribution: The traditional conditional independence test based on vanishing partial correlations assumes that the model variables are (1) multivariate Gaussian distributed and (2) that the relationships between the variables are linear. The approach used here thus only captures the linear relationships between the variables, even though the true physical relationships can have both linear and nonlinear components. Further, using partial correlation as the conditional independence test may not be ideal when a predictand is binary, as discussed in Section 4.2.2.

2) Class imbalance - small number of events per phase: For each of the four causal models presented in Fig. 4.7 (and Fig. 4.8), the samples of the binary MJO variables contain a relatively small number of ones compared to the number of zeros. This is because the number of samples indicating an MJO event occurring in the phase of interest is fairly small. The class imbalance created by this small sample size can reduce the accuracy of the learned graphical model to faithfully represent the true relationships of the population.

3) Latent Confounders: Another important assumption is that the model has no hidden common causes, i.e. latent confounders. In practice, we cannot guarantee that there are no hidden common causes to the variables included in the model. Therefore, we emphasize that the relationships iden-

tified by the causal discovery approach need to be carefully interpreted only as *potential* causal relationships.

For a deeper discussion of these and other assumptions, see the review paper by [102]. It should be emphasized that, in spite of the above limitations, this approach is considerably *more rigorous*, and provides more reliable insights into true causal relationships, than the standard practices of regression and composite analyses currently employed in climate science to identify causal relationships.

We gratefully acknowledge Prof. Clark Glymour and Dr. Joseph Ramsey at Carnegie Mellon University for their feedback on handling mixed binary-continuous variables in causal discovery methods. This research has been conducted as part of the NOAA MAPP S2S Prediction Task Force and supported by NOAA grant NA16OAR4310090. Support was also provided through NSF grant AGS-1445978 (Ebert-Uphoff). All data used in this study is publicly available from the Australian Bureau of Meteorology and the European Centre for Medium-Range Weather Forecasts.

5 Interactions between Synoptic- and Planetary-Scale Atmospheric Disturbances[6]

Atmospheric waves, which are periodic disturbances in the Earth's atmosphere, have varying temporal and spatial scales. These atmospheric waves interact with each other, and how they interact plays a fundamental role in our understanding of climate and weather phenomena. In this case study, we focus on understanding how atmospheric disturbances of different spatial scales interact with each other. We use a spectral decomposition based on spherical harmonics of the daily 500mb geopotential height data field to derive a representation of the different spatial scales. 500mb geopotential height, which is the height at which the air has a pressure of 500mb, is used as it is known to be the steering level for many atmospheric phenomena. We use the PC stable algorithm to hypothesize potential causal interactions.

This initial study focuses specifically on the interactions between two regimes of spatial scales, i.e., the Planetary scale (large scale, e.g., Rossby waves) and Synoptic scale (smaller scale, e.g., tropical cyclones and hurricanes). This analysis acts as a proof of concept study as the interactions between Planetary – Synoptic scales are better known within the climate science community, compared to interactions between other regimes. The preliminary results we present here are encouraging and in-line with the current understanding of atmospheric dynamics. This type of data-driven approach allows getting insights into the causal interactions between specific spatial scales with little effort compared to the analytical methods traditionally used in climate science. Further, this allows studying interactions while taking feedback into account.

[6]This chapter contains material that has been published in the *Journal of the Atmospheric Sciences* as: S. M. Samarasinghe, Y. Deng, and I. Ebert-Uphoff: A Causality-Based View of the Interaction between Synoptic- and Planetary-Scale Atmospheric Disturbances. *Journal of the Atmospheric Sciences*, **77**, 925-941, `https://doi.org/10.1175/JAS-D-18-0163.1`.

In a methods perspective, setting up this case study as a graphical causal model and extracting the relevant information out of the model presented several challenges. As the spectral decomposition used here results in time series of projection coefficients that are complex numbers, we had to think of meaningful ways to represent these time-series in our implementation of the PC stable learning algorithm that can only accommodate real numbers. In this study, we gave particular thought to data preprocessing to ensure that we extract the relevant signals from noise. Also, this is a high dimensional application compared to the two previous case studies. Sections 5.1–5.4 provide details of this case study, while Section 5.5 gives an overview of the case study from a methods perspective.

5.1 Introduction

The Earth's atmosphere is characterized in general by motions of continuous temporal and spatial scales. The interactions among different scales often form the foundations of various weather and climate phenomena. For example, sub-weekly, synoptic-scale disturbances constituting storm tracks are known to play an important role in the development of atmospheric blocks, i.e., quasi-stationary, vertically coherent high pressure features in the extratropical atmosphere (e.g., [167–171]). Initial baroclinic development at long synoptic-scales followed by increasingly important barotropic growth often characterizes the lifecycle of persistent (thus low-frequency) negative height anomalies over the North Pacific in boreal winter that project effectively onto the Pacific-North America (PNA) teleconnection pattern [172]. Lau and Holopainen [173] showed in the quasi-geostrophic (QG) framework that vorticity and heat fluxes associated with both high-frequency (synoptic-scale) and low-frequency eddies act together to maintain the winter monthly mean flow in the Northern Hemisphere. In a broader sense, there exists a symbiotic relationship between extratropical synoptic- and planetary-scale disturbances: the former extract energy from the zonal flow to compensate for their own energy dissipation and supply energy to the latter through barotropic inverse energy cascade; the latter form regions of enhanced baroclinicity where the former preferentially grow [174–176]. As the dominant mode of low-frequency variability in the northern extratropical atmosphere, the Northern Annular Mode (NAM) in the troposphere is char-

acterized by meridional meandering of the jet in the zonal-mean zonal wind and such movement is also largely driven by westerly momentum fluxes of sub-weekly synoptic-scale disturbances of baroclinic origins (e.g., [177–180]).

The investigation of atmospheric scale-interaction processes in the past has relied on a combination of band-pass or spatial filtering of relevant fields with dynamical diagnoses based upon the evaluation of local geopotential tendency, vorticity and/or energy budget (e.g., [173, 174, 181–189]). These approaches are easy to comprehend and when applied to observational or model data often provide excellent depictions of where and how strongly active scale-interactions occur in the physical or frequency domain. One of the biggest drawbacks of the temporal/spatial filtering, as recently argued in a series of papers, is that multiscale energy in physical sense cannot be appropriately defined with just any filters (e.g., [171, 190, 191]). Two minor drawbacks of the filtering are that these approaches do not automatically reveal the exact structure of significantly interacting disturbances given the pre-defined temporal or spatial filtering, and an examination of the temporal evolution of the interaction of interest often demands case-compositing requiring prior knowledge of the occurrences of events, e.g., a vorticity budget analysis applied to multiple blocking events to understand the contribution of synoptic-scale disturbances to the blocking development.

Some research groups have taken an entropy based approach. Liang and Kleeman [192, 193] proposed a rigorous formalism of information transfer (flow) between dynamical system components for both discrete mapping and continuous flow, and the transfer is measured by comparing the entropy increases between an original system under consideration and a modified system where the source component is instantaneously frozen. Liang (2014) derived from first principles (entropy) a measure of causality based on the notion of information flow and used this measure to unravel the cause-effect relation between time series, and specifically, to understand the causal relation between two modes of tropical SST variability (i.e., ENSO and the Indian Ocean Dipole). Liang [194] further provided a comprehensive discussion of the concept of information flow (transfer) and demonstrated information flow and causality as rigorous notions ab initio. Further, the studies in [6, 195] use conditional mutual information in a phase-amplitude domain derived us-

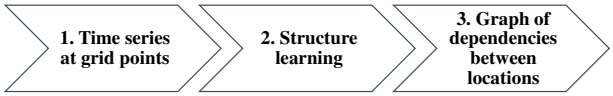
ing the continuous complex wavelet transform to study causal interactions between processes that evolve in different time scales.

In this study we seek to bring a novel approach to understanding scale-interaction processes in the atmosphere, namely the framework of causal discovery and specifically structure learning for temporal probabilistic graphical models (e.g., [20, 28–30, 196, 197]. Probabilistic graphical models (PGMs) have been used in multiple areas of atmospheric sciences. In a first wave of related research PGMs were mainly used for prediction tasks, such as the predictions of severe weather (e.g., [198]), daily pollution levels (e.g., [199]), and precipitation (e.g., [200]). In the second wave structure learning of PGMs and similar tools (e.g., Graphical Granger models) have been used to identify potential cause-effect relationships from data (e.g., [1, 2, 14, 40, 42–44, 83, 201–203]). Runge [102] provides an excellent overview of the practical aspects of constructing PGMs from time series. The authors' earlier work (e.g., [1]) provides to the climate research community an introduction of structure learning using probabilistic graphical models. It shows how structure learning can be used to derive hypotheses of causal relationships among four prominent modes of atmospheric low-frequency variability in the northern extratropics. The idea was further expanded to obtain a climate network that emphasizes flow of information (defined by the directions of directed edges in a PGM) in the 500mb geopotential field and the strength of information flow (in terms of the number of directed edges coming out of one geographical location) in such a network was shown to be decreasing in the future climate with enhanced forcing from greenhouse gases [2, 204]. Note that compared to more rigorous quantitative definitions such as those adopted in Liang [194], the concept of "information flow" is used here more qualitatively. Namely existence of information flow here indicates that the algorithm identifies directions/time delays of statistically significant edges along which information flow occurs, but it does not provide measures of the actual amount of information transferred, i.e. the quantitative strength of the information flow.

The key of structure learning is to detect direct connections and eliminate indirect ones through the use of conditional independence tests [20, 28, 29, 196]. The identified direct connections, either between different teleconnection indices (e.g., [1]), or between geographical locations for a

selected atmospheric variable (e.g., geopotential height, [2]) provide a straightforward view of potential causal pathways in the field of interest. Figure 5.1(a) indicates the process of causal discovery in grid space, as used in the latter type of study. Ebert-Uphoff and Deng [205] tested the performance and accuracy of this approach for dynamic systems using advection-diffusion simulations as a testbed. Others have studied the general algorithms thoroughly for other settings, see for example (Ramsey and Andrews [39]) and many others.

(a)



(b)

Figure 5.1: Basic block diagrams of a) the process of causal discovery in grid space and b) the process of causal discovery in spectral space. Here "SH" stands for Spherical Harmonics.

Here we apply causal discovery to the time series of the daily coefficients of spherical harmonics obtained from a spectral decomposition of the Northern Hemisphere 500mb geopotential field to identify the most prominent direct connections between different spherical harmonics components, and use these potential causal pathways to build a causality-based view of the interaction between atmospheric disturbances of different scales (i.e., spherical harmonics components). Figure 5.1(b) shows the proposed approach of causal discovery in spectral space. There have been a few past and ongoing efforts that study information flow and causality between different scales

by the geophysical fluid dynamics community (most notably [206–209]). However, to the best of our knowledge, only one other research group has combined structure learning with spherical harmonics, namely Zerenner et al. [44]. Zerenner et al. discussed the physical basis why modeling atmosphere as a set of oscillators is more justified in spectral domain compared to in spatial domain: spherical harmonics are associated with large-scale propagating Rossby-Haurwitz waves in the atmosphere. The nature of the connections identified was also elaborated as arising through nonlinear advection of the extratropical quasi-geostrophic flow. However, they only performed a static analysis on monthly mean atmospheric fields, and the insights they obtained using this approach were limited, as their graphs are very sparse with connections primarily between neighboring nodes. Zerenner et al. thus cautioned about deriving a network via thresholding and the potential issue of multivariate Gaussian assumption of the spherical harmonics coefficients. Our results, on the other hand, show many additional connections compared to those reported in Zerenner et al. This different outcome is likely due to several differences between their approach and ours, e.g., we derive a temporal model, use a different data preprocessing scheme, and a different structure learning method. Thus, expanding on the original idea proposed by Zerenner et al., we develop a method that succeeds in identifying many interesting connections in spectral space. The initial focus of this work is on the interaction between synoptic- and planetary-scales which have received most attention in the past. Following this introduction, Section 5.2 describes the data used and detailed analysis steps. The main findings are provided in Section 5.3. Section 5.4 gives some concluding remarks.

5.2 Data and Methods

In this study, we use the daily 500mb geopotential height data at 2.5^{o}by 2.5^{o}horizontal resolution from the NCEP-NCAR reanalysis [210,211] to derive daily fields of atmospheric disturbances of various spatial scales. The study focuses on the Northern Hemisphere and covers all boreal winter months (December-January-February, DJF) in the period 1948 to 2015. The winter months are chosen due to the prominence of eddy-mean flow interaction and scale-interactions in the winter season. To focus on the Northern Hemisphere the time series at locations North of the equator are

mirrored along the equator onto the Southern Hemisphere [212]. This yields the time series at grid points, which is shown in Block 1 of Figure 5.1(b).

The next step is to apply spectral decomposition (Block 2 of Figure 5.1(b)). The daily geopotential height data is decomposed into disturbances of various spatial scales in terms of spherical harmonics, which form a complete set of orthonormal basis functions capable of modeling functions on a sphere (e.g., [212], see Appendix C.1 for review). The properties of the spherical harmonic basis functions characterizing spatial scales are described by the total wave number (L) and the zonal wave number (M). The daily geopotential height data are projected onto these basis functions to obtain daily time series of complex spherical harmonics coefficients (Block 3 of Figure 5.1(b)) as a function of L and M [213]. The coefficients for which $(L + M)$ is an odd number vanish automatically due to the mirroring of the Northern Hemisphere. This leaves only half of the original number of spherical harmonics coefficients for the study.

To apply structure learning of graphical models we first have to define the nodes of the graph (Block 4 of Figure 5.1(b)). To represent the complex coefficient of each spherical harmonics component there are two primary alternatives, namely we can choose the nodes to represent magnitude and phase, or real and imaginary part, of each coefficient. Either alternative results in each spherical harmonics component being represented by two separate nodes in the graphical model. We first tried the magnitude and phase representation, because it appears to be more physically meaningful, as those quantities can be related individually to the power and direction of propagation of atmospheric waves. However, the distribution of the phase variables are cyclic and non-Gaussian, which violates a key assumption of many structure learning algorithms. (Algorithms that do not assume Gaussian distributions exist, but they are of higher computational complexity, and thus generally not feasible for such a large number of variables as required here.) Thus we switched to using the real and imaginary part of the spherical harmonics coefficients instead, which matches the choice by Zerenner et al. [44] whose algorithms also assume Gaussianity. The statistically significant edges in the graphical model identified in the next step among all the nodes (corresponding

to different spherical harmonics) thus represent potential causal interactions among these spherical harmonics, i.e., atmospheric disturbances of different spatial scales.

A customized truncation scheme is adopted to determine how many projection coefficients should be used for structure learning (also Block 4). A triangular truncation with $0 < L \leq L_{max} = 50$ and $0 \leq M \leq L$ is applied first. We then impose an additional constraint that the temporal mean magnitude of the projection coefficient needs to be greater than a specific threshold (0.4 m in our case) to ensure a good signal to noise ratio such that the structure learning techniques perform well. This truncation results in a total of $N = 570$ nodes for a static model (no time lags considered).

Finally, we apply structure learning (Block 5) to the time series of the selected nodes to find potential causal interactions between atmospheric disturbances of different spatial scales. We use a constraint-based structure learning technique that is based on probabilistic graphical models, in contrast to the study by Zerenner et al. that uses a Graphical-LASSO approach. Namely, we use the PC stable algorithm developed by Colombo and Maathuis [23, 38] which is a modification of the classic PC algorithm [37] for improved robustness and speed. (For its use in climate science, see also [1, 2].) The algorithm starts with a fully connected graph, where each node is assumed to have a direct causal interaction with every other node. Then a procedure based on conditional independence tests identifies and eliminates indirect interactions to obtain a final set of direct causal interactions. We want to emphasize here that using this causal discovery approach, we can only detect potential causal interactions, since there can always be latent variables, that is, hidden common causes that were not included in the model. Because of that possibility every relationship found can either be a true causal connection, due to a latent variable or both. All identified relationships must thus be treated as hypotheses of causal relationships, rather than taken as a fact. The only way to find out whether these causal hypotheses represent real causal relations is to check whether the identified cause-effect connections can be interpreted in terms of distinct dynamical and physical processes.

In this study, we use Fisher's Z-test for partial correlation with a significance level of $\alpha = 0.05$ to determine statistically significant edges. Furthermore, to understand the directionality and time

delay of the causal interactions among different nodes, a temporal model is developed using an approach first suggested by Chu and Glymour [40] and applied in grid space in Ebert-Uphoff and Deng [1, 2]. Specifically, we consider $S = 11$ time-slices with neighboring time-slices being 2 days apart. This allows us to consider causal interactions occurring at a maximum time lag of 20 days. The temporal model thus consists of $N \times S = 570 \times 11 = 6,270$ nodes. To find the direction of the edges (i.e., direction of the flow of information), we adopt the temporal constraint that nodes can only influence other nodes at the same or a later time. To handle the initialization problem of temporal Bayesian networks obtained from structure learning, we dropped the first time-slice, leaving 10 of the original time-slices for analysis. This process finally yields the graph of dependencies (Block 6). In this graph, we identify interactions that are consistently repeating over the temporal model. If an edge does not show a repetitive nature and pops up arbitrarily, it may indicate a false discovery, so is not shown.

The results presented in Section 5.3 focus on the directed edges among the nodes with non-zero time lags. In other words, we are seeking causal interactions that occur over a finite time period among disturbances of different spatial scales. The initial effort is devoted to the connection between the Northern Hemisphere planetary-scale and synoptic-scale disturbances as many features of the interactions between the two have been well documented (see discussions in Section 5.1). We define two regimes in the zonal wavenumber space representing respectively the planetary-scale ($0 \leq M \leq 3$) and synoptic-scale disturbances ($6 \leq M \leq 10$). The primary reason for excluding zonal wavenumbers 4 and 5 is that these two wavenumbers sit in between classic planetary-scales and classic synoptic-scales, with the former distinctly tied to large-scale orographic and thermal (e.g., land-sea contrast) forcing and the latter associated mainly with baroclinic instability. These two wavenumbers are often affected by disturbances related to low-frequency variability in the extratropics such as blocking and other "persistent anomalies", which are in turn partly connected to synoptic-scale disturbances. Furthermore, the PC stable algorithm used in this study can only robustly detect linear causal effects due to the use of partial correlations in the conditional independence test. This limitation has been demonstrated by applying the algorithm to

identify the known linear and nonlinear causal effects in the Lorenz model and the relevant results are summarized in the Appendix C.2. Given such limitations, we choose to focus the discussion on two groups of disturbances with distinct physics and sufficiently separated in scales (to create a "quasi-linear" condition) by excluding zonal wavenumbers 4 and 5 in all the figures even though these two components are included in the actual process of structure learning.

As our graphical model has separate nodes for the Real and Imaginary components of each spherical harmonic, we need a suitable mapping to the (M, L) space to be able to interpret these results. Based on our mapping, for the two defined regimes, when one or more of the four conditions listed below are met, we claim that a node in Regime I influences (causes changes in) a node of Regime II:

1. Regime I real part node $\rightarrow$ Regime II real part node,
2. Regime I real part node $\rightarrow$ Regime II imaginary part node,
3. Regime I imaginary part node $\rightarrow$ Regime II real part node,
4. Regime I imaginary part node $\rightarrow$ Regime II imaginary part node.

This definition – with the roles of Regime I and Regime II reversed - also applies to the situation where a Regime II node influences (causes changes in) a Regime I node. To obtain the underlying structure of the pair of (potentially) interacting disturbances from the two regimes, we also re-construct daily geopotential height fields for the two regimes using spherical harmonic components in one regime that are found to be actively interacting with spherical harmonics in the other regime. In other words, a spatially-filtered version of the daily geopotential height field is produced by retaining in the daily height field only spherical harmonics components that are connected by directed edges in Figure 5.2. These reconstructed disturbances can be created for interactions occurring at different time lags when only edges of desired time lags are included in the reconstruction. When considering all statistically significant edges with a time lag equal to 2 days, we get an overall view of a pair of synoptic-scale and planetary-scale disturbances that most actively interact with each other in boreal winter. The choice of two days is arbitrary here. However, given the typical lifecycle of synoptic-scale disturbances (less than a week), we expect

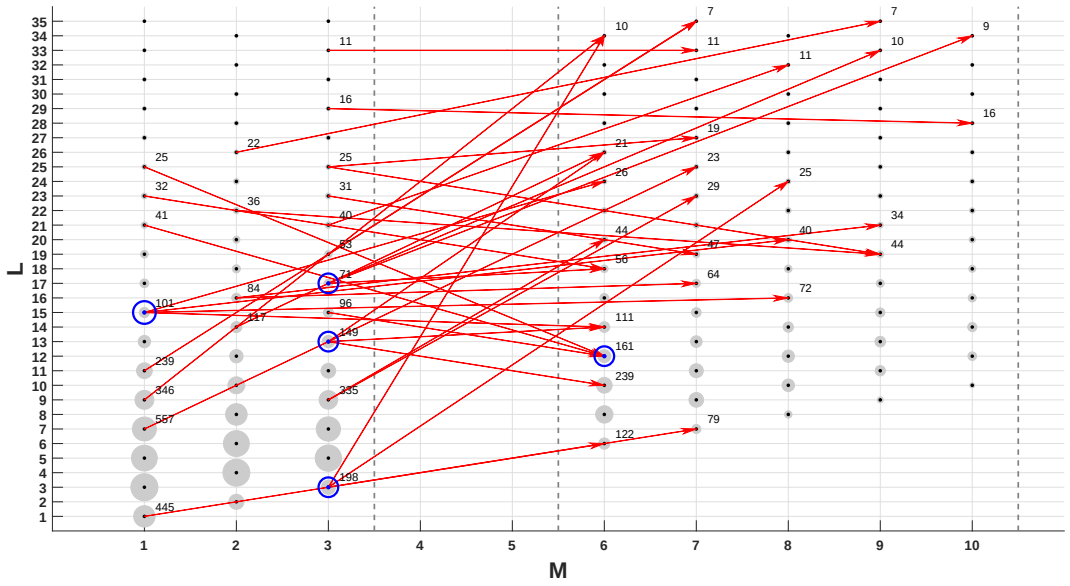
the modulation/feedback effects between the two regimes to occur at a rather short timescale (2-3 days).

5.3 Results

Figure 5.2 provides an overview of the interaction between planetary- and synoptic-scale disturbances in Regime I and Regime II, respectively. A directed edge (a red arrow in Figure 5.2) indicates that changes in the disturbance with wavenumbers (M, L) at the beginning of the edge (arrow) tends to cause changes in the disturbance with wavenumbers at the end of the edge (arrow). In other words, the directed edge depicts the direction of the flow of information in the wavenumber space. The presence of a blue circle around a node means that there are more than 2 incoming (or outgoing) edges toward (or out) of this node, suggesting a relatively active role played by this node in the interaction between Regime I and Regime II disturbances. The size of the blue circle is proportional to the number of edges associated with the node. The size of the gray shaded circle around each node indicates the winter mean magnitude of the corresponding spherical harmonics component. Based on these magnitudes, it is evident that the identified causal structure is not only driven by the magnitudes of the coefficients. Further, to reduce the impact of false discoveries, we do not show edges that only occur once in the temporal model. This information, viewed together with the directed edges of the node, provides a more concrete idea how physically important an identified connection is, as a connection of statistical significance does not always mean that the connection is "substantial" and the number of edges is not equivalent to the magnitude of the flow of information.

Figure 5.2a shows all edges pointing from Regime I to Regime II nodes, representing the influence of planetary-scale disturbances on synoptic-scale disturbances. The total number of such edges is large (≈ 80). However, the energy of disturbances with large meridional wavenumbers ($L - M > 6$) is rather small [174] and this naturally places more weights on the edges identified at the lower-right corner of the (M, L) parameter space in Figure 5.2b. Judged by the size of both the blue circle and gray shaded circle, synoptic-scale disturbances $(6, 10)$, $(6, 12)$, $(6, 14)$, $(7, 17)$ and $(8, 16)$ are under the most obvious influence of planetary-scale disturbances

(a)



(b)

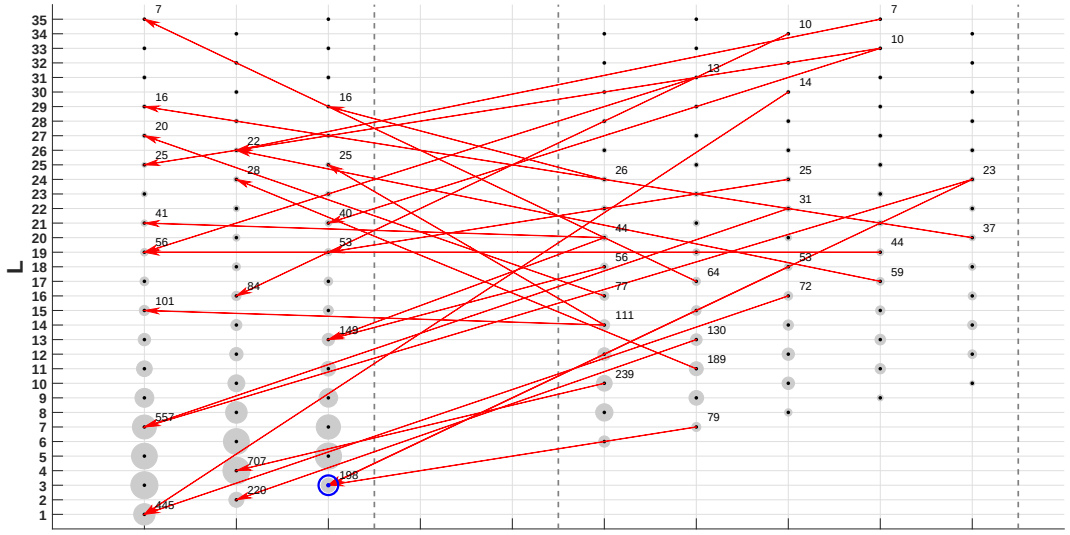


Figure 5.2: Directed edges between Regimes I and II with respect to the total wave number (L) and the zonal wave number (M). a) Regime I to II b) Regime II to I. Black dots denote the nodes used for the analysis. A blue circle around a node indicates that the node has more than 2 incoming or outgoing edges. The size of each blue circle is proportional to the number of edges associated with the node. The gray shaded circles are proportional to the winter mean magnitude of the coefficient of the corresponding spherical harmonics component. The number alongside each interacting node denotes the winter mean magnitude of the spherical harmonics coefficient scaled by 10.

with zonal wavenumbers ranging from 1 to 3. Figure 5.2b displays the edges characterizing the impact of synoptic-scale on planetary-scale disturbances (Regime II to Regime I). Disturbances $(6, 10)$, $(6, 14)$, $(6, 16)$, $(7, 11)$, $(7, 13)$ and $(8, 16)$ stand out as those providing most active feedback to planetary-scale disturbances. Disturbance $(2, 4)$ has the largest mean magnitude among all planetary-scale disturbances influenced by feedbacks from synoptic-scale disturbances and this verifies the traditional picture that zonal wavenumber 2 planetary wave actively interacts with the two Northern Hemisphere storm tracks in boreal winter (e.g., [214, 215]).

The reconstructed daily 500mb geopotential height field using only spectral components that are interacting according to Figure 5.2 provides a direct view of the spatial structure of the actively interacting planetary-scale and synoptic-scale disturbances. Once the interacting spectral components are identified, the specific disturbance can be reconstructed as a linear combination of the spectral components (spherical harmonics bases). For a specific reconstruction, we use each identified spectral component only once, regardless of the number of interactions it has with other spectral components. Figure 5.3 displays the total (Figure 5.3a) as well as the reconstructed geopotential height field for Feb 16, 2001. The middle two panels (Figures 5.3b and 5.3c) are respectively the Regime I disturbance and the Regime II disturbance that is being actively influenced by the Regime I disturbance on Feb 16, 2001. In the extratropics, the planetary-scale flow is characterized by a roughly zonal wavenumber-1 structure with a major trough (negative geopotential height anomalies) extending from the east coasts of Asia toward North America and a major ridge (positive geopotential height anomalies) over Western Europe (Figure 5.3b). This wavenumber-1 planetary-scale disturbance modulates the activity of synoptic-scale disturbances shown in Figure 5.3c, which manifest themselves as a classic circumglobal wave train consisting of meridionally elongated disturbances. The amplitude of these shortwave disturbances peaks over the North Pacific and North Atlantic, indicating significant baroclinic growth and subsequent downstream development over these two ocean basins and thus the presence of the two major storm tracks (e.g., [216]). The Regime I to II disturbance structure revealed here is consistent with the classic theory of storm track dynamics where storm track disturbances (cyclones/anticyclones)

preferentially develop downstream of the planetary-scale troughs (and time-mean jets) as a result of locally enhanced baroclinicity at the location of the jets and further downstream development associated with ageostrophic flux of geopotential (e.g., [174, 216]).

Figure 5.3d and 3e depict the influence of synoptic-scale disturbance on the planetary-scale disturbance. Specifically, the synoptic-scale disturbances in Figure 5.3d are producing a zonal wavenumber-2 structure with negative height anomalies off the east coasts of Asia and North America and positive height anomalies over the eastern North Pacific and Western Europe. This wavenumber-2 structure locally enhances the trough off the east coast of Asia and the ridge over Western Europe as shown in Figure 5.3b. Therefore, the Regime I disturbance is responsible for the excitation/propagation of the Regime II disturbance whose dynamical feedback to the Regime I disturbance locally re-enforces the trough and ridge pattern in the original wavenumber-1 structure of the Regime I disturbance. This result again verifies the notion that winter planetary-scale flow and the associated zonal wind jets tend to be "self-sustained". The feedback of synoptic-scale to planetary-scale disturbances is achieved through the vorticity (momentum) and heat flux by synoptic-scale disturbances (e.g., Lau 1984). In Figure 5.3d, synoptic-scale disturbances change the alignment of their major axis from being SW-NE (Southwest-Northeast) oriented over the central North Pacific to being NW-SE (Northwest-Southeast) oriented over western North America. The associated meridional flux of zonal momentum by these disturbances subsequently changes from a poleward to an equatorward direction. The poleward (equatorward) flux of zonal momentum is consistent with a poleward (equatorward) shift of an eddy-driven westerly jet and thus the formation of a negative (positive) height anomaly on the poleward side of the disturbance over the central North Pacific (western North America) following the argument of geostrophic balance (Figure 5.3e). The same mechanism also applies to the North Atlantic and Western Europe, responsible for the formation of the negative and positive height anomalies there, respectively. Comparing the reconstructed height fields with the total height field (Figure 5.3a), we note that the planetary-scale disturbances depicted in Figs. 3b and 3e miss another major trough over central Eurasia. This is

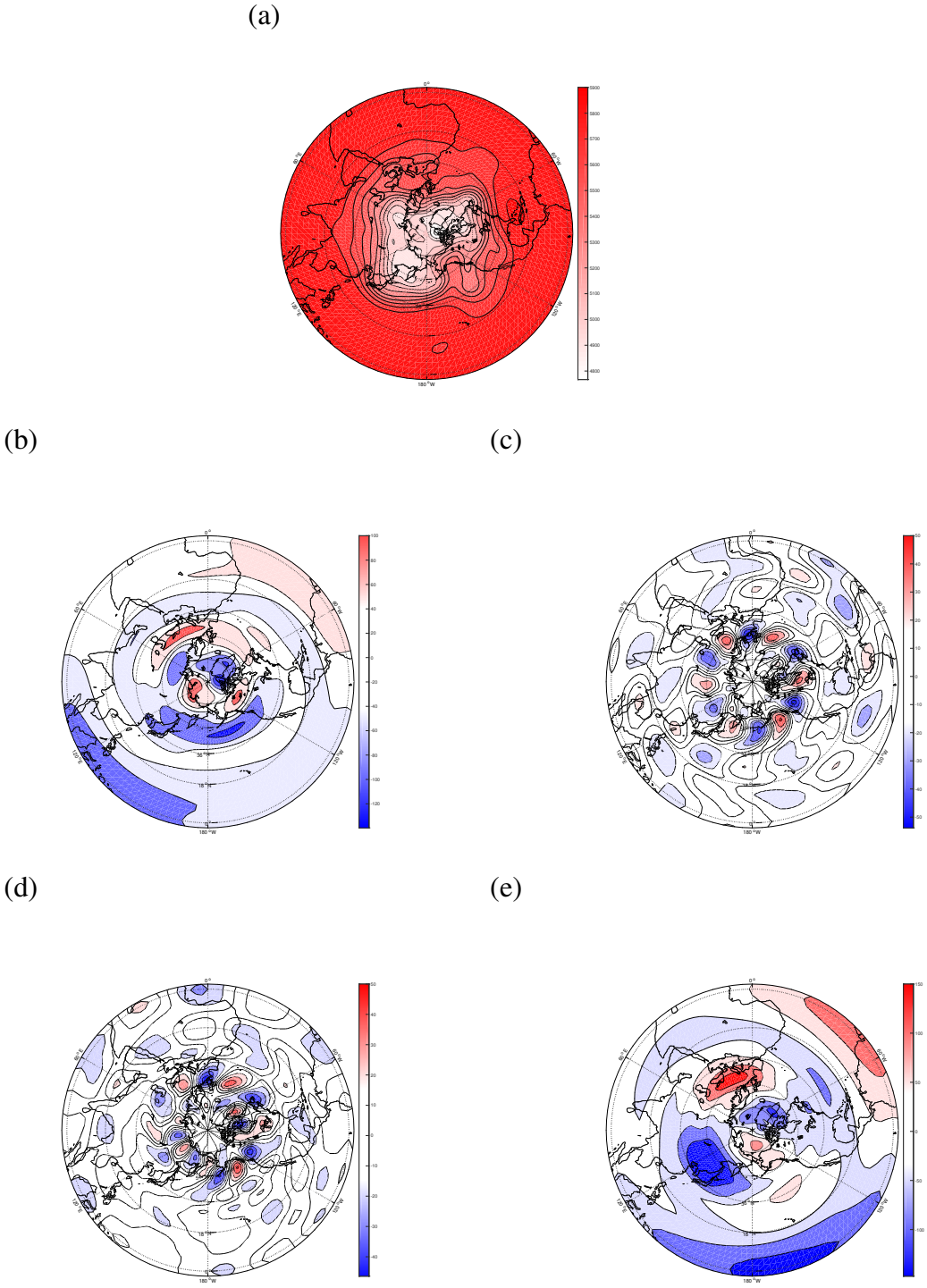


Figure 5.3: Reconstruction of the daily geopotential height using only the significantly interacting spectral components and excluding the interactions with M=0. Upper panel: The 500mb geopotential height field on Feb 16, 2001 (a). Middle panel: Interacting disturbances from Regime I (b) to Regime II (c) (Feb 16, 2001). Lower panel: Interacting disturbances from Regime II (d) to Regime I (e). (Feb 16, 2001). Unit: m.

likely the result of excluding zonal wavenumbers 4 and 5 from the figures, for reasons described in Section 5.2.

Figure 5.4 shows the winter climatological variance (root-mean-square (RMS)) of the daily geopotential height of the actively-interacting disturbances identified in Figure 5.2 and presented in Figure 5.3. When the flow of information from Regime I to Regime II (i.e., planetary-scale disturbances influencing synoptic-scale disturbances) is being considered, the dominant day-to-day variability in the planetary-scale flow peaks at the trough and ridge locations of the zonal wavenumber-1 disturbance (as shown in Figure 5.3b). The corresponding variability in synoptic-scale flow maximizes along a well-defined circumglobal band in the extratropics with elevated amplitudes over the North Pacific and North Atlantic, where the two climatological Northern Hemisphere storm tracks reside (Figure 5.4b). In the opposite direction of information flow, when synoptic-scale disturbance affects planetary-scale disturbances, storm tracks again characterize the day-to-day variability in the synoptic-scale flow (Figure 5.4c) while the variability in the planetary-scale disturbance reflects the zonal wavenumber-2 structure that is effectively forced by synoptic-scale disturbances and locally enhances the trough and ridge of the zonal wavenumber-1 structure of the winter planetary-scale flow shown in Figure 5.3b (Figure 5.4d).

To see whether there is any systematic change in the amplitude of the actively-interacting planetary- and synoptic-scale disturbances, we plot in Figure 5.5 the seasonal mean amplitude of the disturbances as a function of time. The amplitudes of the daily reconstructed disturbances are spatially averaged over the sphere and temporally averaged over each winter season to create these time series. Blue and orange curves in Figure 5.5a are respectively the mean amplitude of Regime I and Regime II disturbances when Regime I is influencing Regime II. No significant trends can be identified in these time series, but substantial variations occur across interannual to interdecadal time scales. Specifically, mid-1970s to mid-1980s are characterized by above normal amplitudes of disturbances when Regime I influence on Regime II is being considered. Similar results are found in the case of the feedback of Regime II disturbance to Regime I disturbance (Fig. 5.5b). There appear no obvious connections between the activity levels of these wave-wave

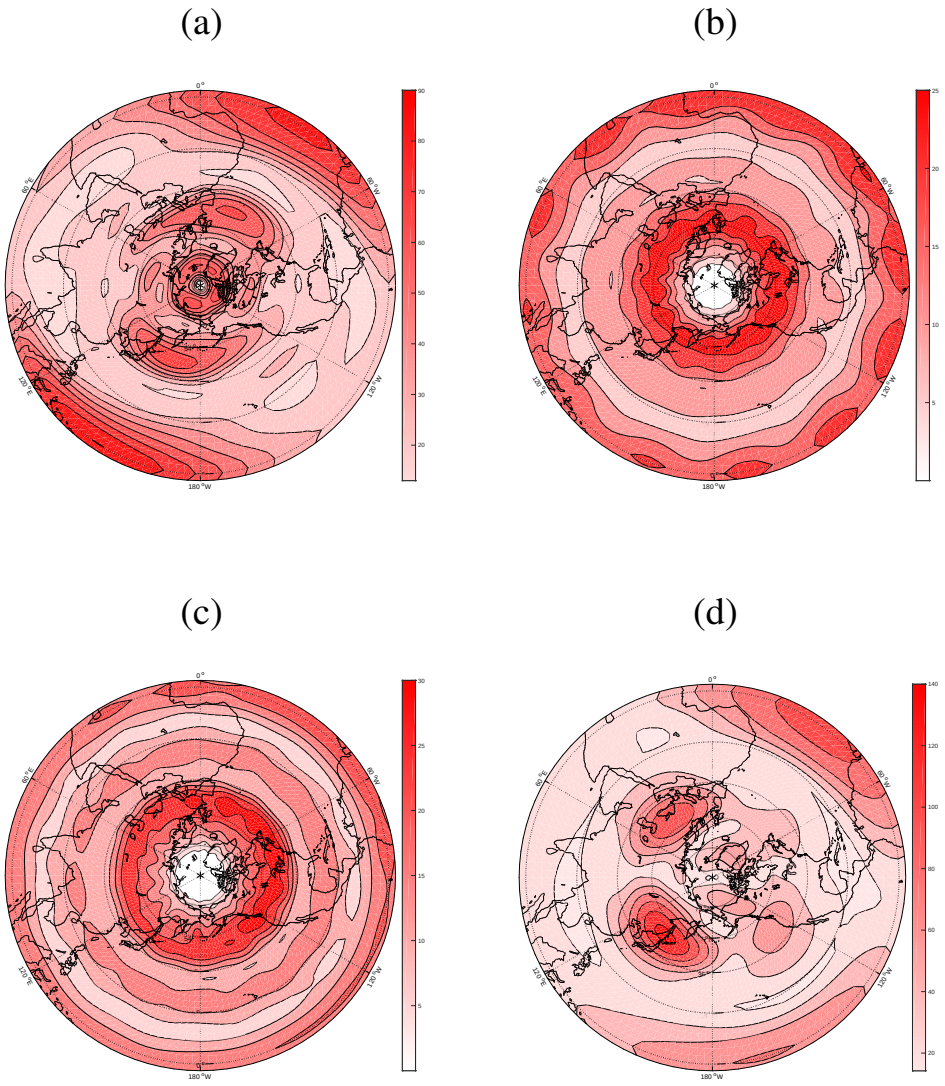


Figure 5.4: Variance (RMS) of the geopotential height reconstructions of the significantly-interacting disturbances, excluding the interactions with M=0. Regime I (a) to II (b); Regime II (c) to I (d). Unit: m2.

interactions and increasing greenhouse gas (GHGs) forcing, nor between the activity levels and primary modes of variability in the climate system, including the ENSO and the Pacific Decadal Oscillation (PDO). These results suggest that the interactions identified by our approach are likely features internal to the atmosphere and the low-frequency components of the time series shown in Figure 5.5 are mainly the result of nonlinearity in the system.

5.4 Concluding Remarks

This paper reports some preliminary yet encouraging results concerning the use of constraint-based structure learning to understand scale-interaction processes in the atmosphere. The analysis focuses on identifying causal pathways among atmospheric disturbances of different spatial scales. Temporal probabilistic graphical models illustrating such causal pathways are built by applying the PC stable algorithm to the spherical harmonics decomposition of the boreal winter daily 500mb geopotential height data during the period 1948-2015. With an initial focus on interplays between planetary-scale (Regime I) and synoptic-scale (Regime II) disturbances, the identified directed edges (information pathways) suggest active coupling between Regime I disturbances with zonal wavenumber-1 to wavenumber-3 structures and Regime II disturbances with zonal wavenumber-6 to wavenumber-8 structures. These scales fall nicely into the classic picture depicting interactions between Northern Hemisphere planetary waves and storm tracks in boreal winter.

Furthermore, daily reconstruction of geopotential heights using major nodes (in the (M, L) wavenumber space) connected by the detected causal pathways suggest that the modulation of synoptic-scale disturbances by planetary-scale disturbances in the northern extratropics is best characterized by the flow of information from a zonal wavenumber-1 disturbance to a synoptic-scale circumglobal wave train whose amplitude peaks at the North Pacific and North Atlantic storm track region. The feedback of synoptic-scale disturbances to the planetary-scale disturbances manifest itself as a zonal wavenumber-2 structure driven by synoptic-eddy momentum fluxes that locally enhances the East Asian trough and Western Europe ridge of the original wavenumber-1 structure in the planetary-scale disturbances that are actively modulating the activity of the synoptic-scale disturbances. The seasonal mean amplitude of the significantly-interacting disturbances detected

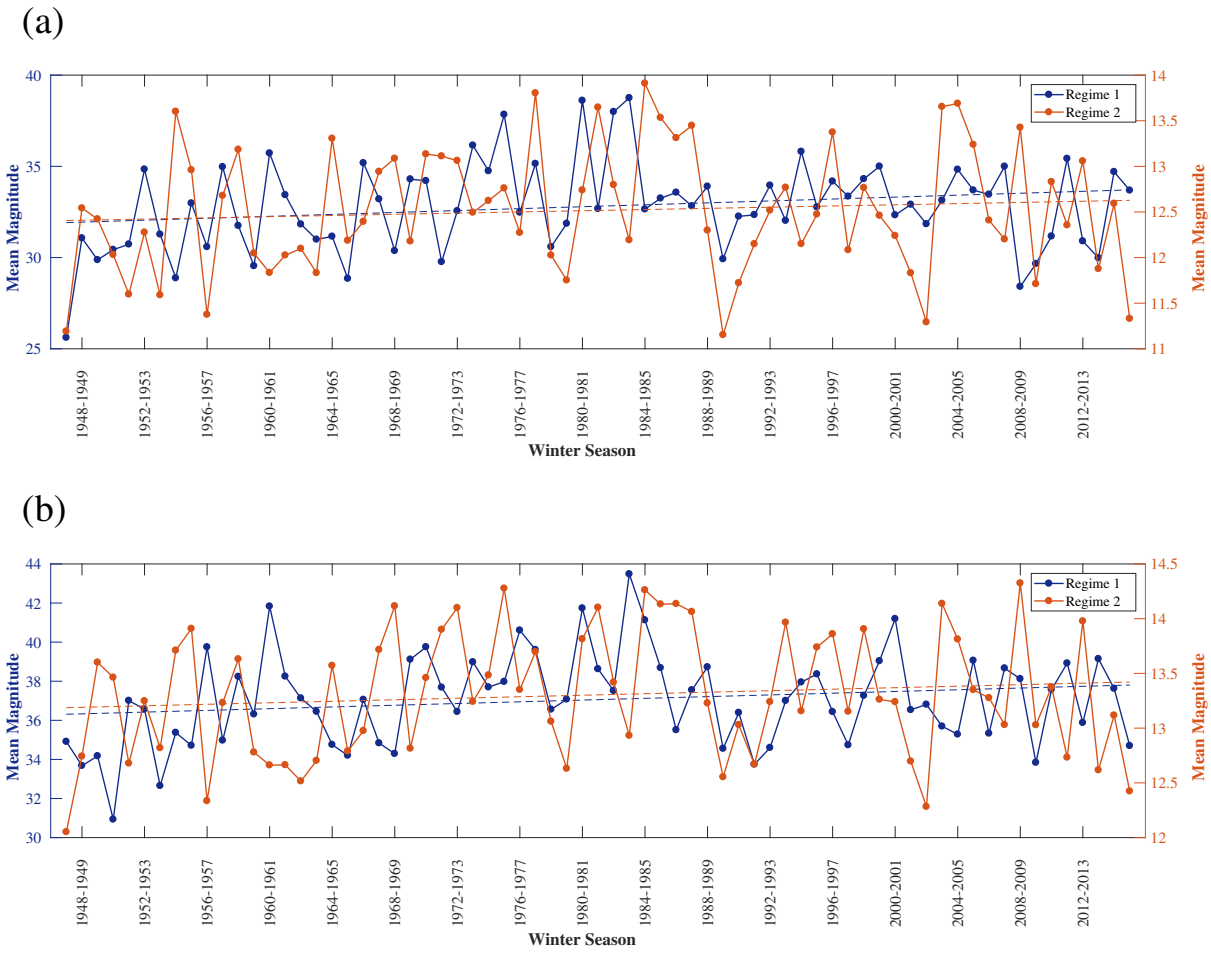


Figure 5.5: Annual mean magnitude time series of interacting disturbances excluding M = 0 interactions. a) Regime I-II interactions; b) Regime II-I interactions. A spatially averaged magnitude over the sphere is temporally averaged for each winter season from 1948 to 2015. Regime I and II time series are in blue and orange, respectively. The blue/orange dashed lines indicate the corresponding linear trends. Unit: m.

here exhibits pronounced variations across interannual-to-decadal time scales that are not correlated with major low-frequency modes of variability such as the PDO in the climate system.

The analysis presented here provides a new way to examine the scale-interaction processes in the atmosphere in the context of causal discovery and structure learning for probabilistic graphical models. The PC stable algorithm can be applied to identify potential causal pathways among atmospheric disturbances of various spatial scales. The structure of actively-interacting pairs of disturbances can be reconstructed from nodes in the graphical models that are connected by statistically significant edges, making it easier to describe the temporospatial characteristics of interacting scales in the atmosphere. New metrics based on the detected edges such as the mean amplitude of the interacting disturbances can be computed for model simulations and observations to add a new dimension to the validation of dynamical properties of simulated atmospheric variability. Despite all the benefits of the approach, the application of the PC stable algorithm to real atmospheric and climate data must be pursued on a cautious note due to the assumed linearity of interactions in the conditional independence test. The scale-interaction in the real atmosphere is largely realized via nonlinear advection processes. Therefore, in our study the cause-effect relations (i.e., directed edges) among spherical harmonics components identified by the PC algorithm are most robust and physically relevant only when they are between spherical harmonics components that are sufficiently different in scale. The scale-separation ensures one of the components remains relatively "temporally steady" and/or "spatially uniform" during the interaction and creates a "quasi-linear" condition for the algorithm to be applicable. The implication of this limitation is that edges identified between neighboring spherical harmonics components in our analysis are subject to much greater uncertainties. Our aggregation of results into different regimes helps to create this 'quasi-linear" condition and excluding the scales at the Regime I and II interface (i.e., zonal wavenumbers 4 and 5) from the discussion serves to further reduce uncertainty in the results presented. However, the difficulty in reliably detecting nonlinear causal relationships places a major constraint on this approach's capability of gaining a significant amount of new insights into the cross-scale interactions in the real atmosphere. Other approaches designed to identify nonlin-

ear causal effects are being developed (e.g., convergent cross mapping, [217] and causal network learning algorithms that use nonlinear conditional independence tests, e.g., [61])), but the associated computational complexity makes it hard to apply to high dimensional problems such as the one studied here. Another challenge we face in this approach of causal discovery is constructing a quantitative measure of the discovered causality. Compared to the entropy-based information flow, information flow described here does not carry a magnitude associated with known physical quantities. This prevents us from making a direct comparison with the conclusion drawn in [209] emphasizing "bottom-up causation" for mid-latitude atmosphere in a quasi-geostrophic setting. Palus [195] uses conditional mutual information to obtain information flow and the results based on an application to daily surface air temperature data over Europe suggest transfer of information from larger to smaller timescales, different with the "bottom-up causation" found in [209]. All of these highlight the significant uncertainties in causal discovery for atmospheric processes, especially when real observations are being analyzed.

Our ongoing work includes an extension of the analysis to consider the interaction between disturbances in the synoptic- and meso-scale regimes and also to consider the interactions in boreal summer. Finally, note that the advective nonlinearities dictate that scale interactions in a fluid such as the atmosphere occur in the form of wave triads, whose wavenumber vectors must sum to zero. If zonal and meridional scales are quantified in terms of Fourier coefficients and these coefficients serve as inputs for the PC algorithm, any two or three directed edges involving three nodes should have the three corresponding wavenumber vectors (zonal and meridional wavenumber as the two components) sum to zero. However, in spherical harmonics expansion, zonal wavenumbers are integers and meridional scales are quantified through the number of nodes between the poles. The triad interaction therefore does not have a straightforward reflection in the detected edges that span a spectral space defined by the integer zonal and total wavenumber. Future work will include an investigation of the connection between the constructed graphs and wave triads potentially utilizing a Fourier expansion of the daily height field.

5.5 Additional Discussions with a Methods Perspective

The goal of this study is to understand how atmospheric disturbances of different spatial scales are interacting with each other. As discussed earlier, this study poses this research problem as a question of identifying how the time-series of the non-zero projection coefficients (of the relevant spherical harmonic bases) are interacting with each other. In this section, we discuss the key issues and considerations that needed our attention through the process of deriving a graphical causal model from data to extracting useful information out of this model to answer the relevant research questions.

Creating a graphical model in an abstract variable space: The first step of applying structure learning is to define the nodes of the graph. However, as the projection coefficients associated with spherical harmonic bases are complex numbers, we consider two representations. That is, (1) to represent the amplitude and phase components of the coefficient as two separate nodes or (2) to represent the real and imaginary components. Even though an amplitude and phase representation, which is closely related to the power and direction of propagation of atmospheric waves, seems to be the more physically meaningful representation, some practical issues make this path infeasible. First, the phase variable, which is cyclic by nature, does not follow a Gaussian distribution making it infeasible to use the traditional conditional independence test that relies on partial correlation. Second, even though there are conditional independence tests that do not use the assumptions of normality or linearity, these methods are computationally expensive for high dimensional applications. Therefore, we choose to represent the real and imaginary parts of the complex coefficient as two separate nodes in the graphical model following Zerenner et al.

Extracting strong causal signals: The study presented here builds on the work performed by Zerenner et al., where they derive graphs for spherical harmonics based on partial correlation and use the GLASSO algorithm for structure learning. Their approach yields very sparse graphs with edges foremost between first and second-order neighbors in spectral space. As they point out themselves, that is surprising as one would expect dependencies among many more spherical

harmonics. In contrast to their study, we follow specific steps to facilitate the learning algorithm to capture the relevant causal signals.

First, we use daily data based on the recommendation of the domain expert, compared to monthly data used by Zerenner et al. When working with applications in the climate, it is essential to use data with an appropriate temporal resolution. Also, we restrict the study to dates relating to the northern hemisphere winter (Dec-Jan-Feb), i.e., when we expect the relevant causal signals to be prominent. As explained earlier, we mirror the data of the northern hemisphere along the equator to ensure that we only use the signals relevant to the northern hemisphere winter for this study. Further, when selecting the variable in the structure learning method, we use a truncation scheme based on the temporal mean of the magnitude. The truncation scheme reduces the dimensionality of the problem while allowing the causal signals to distinguish from noise.

Extracting relevant information to answer the research questions of interest: The graphical model identifies interactions between Real and Imaginary components of the projection coefficient time-series. However, we are interested in how a projection coefficient time-series interacts with another (not just the Real or Imaginary parts). Therefore, we use a suitable mapping from the (Real,Imaginary) to the (M,L) space to allow physical interpretation of the results. Does the graphical model answer the research questions of interest? If not, how can we extract the relevant information out of the graph? These are questions that we need to consider when working with practical applications.

High dimensionality: The case study presented here is high dimensional with over 6,000 nodes in the graphical model. We implemented an optimized code that parallelizes the conditional independence tests to facilitate structure learning at higher dimensions.

We gratefully acknowledge Clark Glymour & Joseph Ramsey at CMU for constructive discussions related to this project. We also thank three anonymous reviewers for their thoughtful comments and suggestions that led to major improvement of the manuscript. The NCEP-NCAR reanalysis data used in this study was provided through the NOAA Climate Diagnostics Center. This research was supported by the NSF Climate & Large-Scale Dynamics (CLD) program under Grants AGS-1354402 & AGS-1445956 (Deng) & AGS-1445978 (Ebert-Uphoff).

6 Conclusions

This chapter concludes the dissertation by providing a summary of key contributions and by discussing suggested future research directions.

6.1 Summary of Key Contributions

The three case studies contained in this dissertation cover a wide range of questions that are of interest to the climate science community, namely the Arctic mid-latitude connections, teleconnection pathways between the MJO and NAO, as well as the causal interactions between Planetary and Synoptic atmospheric disturbances. We identify key relationships between climate variables as well as the timescales at which the interactions are occurring. As an observational study cannot prove causal interactions, due to reasons such as latent variables, we view the results derived from these studies as *data-driven hypotheses* of causal interactions.

In this research, we use constraint-based structure learning, as well as graphical Granger methods to derive these data-driven hypotheses. These methods allow us to distinguish direct interactions from indirect ones and also identify the overall causal structure in a multivariate setting, providing an advancement from the correlation and traditional lagged regression approaches that are commonly used in climate science. Further, these approaches allow us to get a two-way feedback perspective of these interactions allowing a climate scientist to gain insights that go beyond to what can be obtained from a targeted modeling approach. As discussed in detail in chapters 3 – 5, each of these case studies generates new scientific insights. We do not have a mechanism to determine the accuracy of these results as there is no ground truth for the actual causal interactions between the variables used in these studies. However, through close collaboration with climate scientists, we have been able to validate that these results agree with the current understanding of climate physics and dynamics, and provide new hypotheses that may be evaluated using targeted modeling studies.

Also, we have discussed each case study with a method perspective as the main objective of our research is to promote and make modern causal inference methods more accessible to the climate and Earth science communities. We have documented the scientific thought process of setting up each case study, the challenges faced, and how we have dealt with these challenges. We hope that our discussions will facilitate a researcher/climate scientist who is at entry-level to spatiotemporal causality by providing ideas and suggestions to tackle more general questions in climate science.

The first case study exploring the interactions between Arctic temperature and midlatitude jet-streams compares the results of three causal inference methods. Namely, VAR-Granger and LASSO-Granger, that fall under the Granger causality framework and the PC stable algorithm that falls under the Pearl causality framework. The results show that Arctic temperature is linked together with jet speed and jet position by two robust positive feedback loops. This study shows that the results of all three methods agree very well in low dimensional applications when we use a conditional independence test based on partial correlations with PC stable.

This similarity of results is not surprising as all of these methods use partial correlations as the underlying indicator of relationships. However, the results are not identical because the VAR-Granger models generate partial correlations by conditioning on all the variables simultaneously. PC performs partial correlation tests consecutively, starting from small conditioning sets and going to higher dimensions only as needed. The study shows how VAR-Granger and LASSO-Granger approaches allow a researcher to easily identify both the strength and the timescales of the interactions. In contrast, PC only allows identifying the timescales of interactions. Even though the results of the three methods agree well in low dimensional settings, studies have shown that PC is the preferred choice over VAR-Granger in higher dimensional applications [102]. In this case study, all of these methods identify the average, linear interactions between the variables. However, we have the flexibility to also use PC stable with other appropriate conditional independence tests that will allow us to identify nonlinear interactions.

The second case study included in this dissertation explores the causal pathways of the teleconnections between the Madden Julian Oscillation (MJO) and the North Atlantic Oscillation (NAO).

This research suggests that the MJO can influence the NAO not only through wave propagation in the troposphere but also through a pathway involving the stratospheric polar vortex. In this study, we use climate indices calculated using variations of the principal component analysis to represent the MJO and NAO oscillations. As discussed in detail in Chapter 3, we derive the amplitude and phase of the MJO and then use a binary indicator variable to indicate the occurrence of an MJO event. Identifying physically meaningful nodes/variables for the graphical causal model and setting up the model in a way such that we can use these models to answer the science questions of interest are two of the main challenges/considerations associated with this study. This study also elucidates that, in some scenarios, the graphical model alone will not be sufficient. These methods may need to be used alongside other tools to be able to tackle the more general questions in science.

With the third case study, we present a complex high-dimensional causal inference application and the first successful application of causal inference in spherical harmonic space. This study uses a spectral decomposition based on spherical harmonics along with structure learning based on the PC-Stable algorithm to hypothesize cause-effect relationships between atmospheric waves of different spatial scales. This study builds on the ideas by Zerenner et al. [44] and progresses it to tackle an interesting, yet difficult question in climate science. This study highlights the importance of selecting data with an appropriate temporal resolution for the study and emphasis on preprocessing data such that the inference methods can easily extract the causal signals of interest. This case study tackles challenges related to an abstract variable space associated with the time-series of complex projection coefficients of the spherical harmonics. Also, it showcases the efforts followed in our research to extract relevant information out of the graphical model. We can think of this case study as a proof of concept study as the Planetary-Synoptic scale interactions are relatively well known within the climate science community. Extending this study to different spatial regimes, for example, to study Synoptic-Meso scale interactions, is thought to be able to provide a significant contribution to the climate science community.

In addition to these case studies, we make an effort to improve the accessibility to the causal inference methods by creating two tutorials containing Python-based implementations and step-by-step explanations of 1) the PC-stable algorithm for time-series and 2) the bivariate Granger causality approach. We have made these tutorials publicly available [24, 25].

6.2 Way Forward

The case studies presented in this dissertation attempt to tackle several important questions, challenges, and considerations related to causal inference in climate. However, this is only a small fraction of what a researcher/climate scientist will have to face practically. The case studies presented here all employ linear causal inference techniques. However, we know that the complex interactions between climate variables can be nonlinear and chaotic. We show an example of the limitations of the conditional independence test based on partial correlations with our sample application to the Lorenz system, which is a simplified mathematical model for atmospheric convection (see Appendix C.2). A partial correlation-based method may not capture certain important interactions between variables if the effects cancel out on average. For example, if two variables follow a relationship such that they are correlated for a certain time period then anticorrelated for another, then a correlation-based method may say that these two variables exhibit zero correlation on average through the full duration (see [217] for details).

With the ubiquitous nonlinearity in the atmosphere, it is important to identify methods that can capture nonlinear relationships between variables. Runge et al. provide a version of the PC algorithm along with different conditional independence tests in their TIGRAMITE [61] package to identify both linear and nonlinear causal relationships between time series variables. This package provides a linear conditional independence test based on partial correlation as well as several nonlinear conditional independence tests including a test based on Gaussian process regression and a distance correlation, and a nonparametric test based on the k-nearest neighbor estimator of conditional mutual information [61]. This package has already found promising use in test cases as well as applications related to climate science [3, 4, 61], and we encourage researchers to explore

it. Making these methods more scalable to higher dimensional applications can provide useful insights into questions such as the one presented in Chapter 5.

A limitation of most standard causal inference frameworks is that they assume that no hidden common causes are acting upon the variables included in the model. This assumption is often violated in climate applications resulting in these methods producing spurious/incorrect results. There exist causal inference methods that do not rely on the assumption of causal sufficiency. Identifying the feasibility of these algorithms for the spatiotemporal applications in climate science will provide a significant advancement to causal inference in climate and is a suggested next step to this research. On the Granger causality side, there have been some first efforts to infer causal structure in the presence of hidden variables in climate settings [43, 218]. On the Pearl causality side, algorithms for inferring hidden common causes exist, but to the best of our knowledge, none of them have been adequately tested for climate applications. Primary algorithms are the Fast Causal Inference (FCI) algorithm [19], [20], and tsFCI [21], which is an extension of FCI for time-series data that enforces that the identified causal structure is time-invariant.

Through the case studies presented in this dissertation, it is clear that causal inference applications in climate science require substantial attention to steps related to variable identification as well as setting up the inference problem in a physically meaningful way. These steps need domain expertise of a climate scientist. Also, the final steps of evaluating, validating, and interpreting the results require substantial domain knowledge. Therefore, we highlight the importance of carrying out these research activities as close collaborations between climate and data science communities.

Furthermore, we encourage the research communities focused on method development to make modern causal inference methods more accessible to their end-users. For example, providing code implementations that are publicly available and providing documentation that explains the inference approaches in an easy to understand language can help a climate scientist to get started on their applications with ease. Also, presenting results of how these methods performed for certain test cases can help the user evaluate and understand the usefulness of these methods. The accuracy of these methods, as well as the sensitivity of these methods to method-specific parameter selection

and sample size, can give useful insights that can guide method selection. Finally, providing case studies, like the ones provided in this dissertation, can help a user identify the scope and utility of these methods as well as how these methods can be useful in practical applications.

Bibliography

[1] I. Ebert-Uphoff and Y. Deng, "Causal discovery for climate research using graphical models," *Journal of Climate*, vol. 25, no. 17, pp. 5648–5665, 2012.

[2] I. Ebert-Uphoff and Y. Deng, "A new type of climate network based on probabilistic graphical models: Results of boreal winter versus summer," *Geophysical Research Letters*, vol. 39, no. 19, 2012.

[3] J. Runge, V. Petoukhov, J. F. Donges, J. Hlinka, N. Jajcay, M. Vejmelka, D. Hartman, N. Marwan, M. Paluš, and J. Kurths, "Identifying causal gateways and mediators in complex spatio-temporal systems," *Nature communications*, vol. 6, 2015.

[4] M. Kretschmer, D. Coumou, J. F. Donges, and J. Runge, "Using causal effect networks to analyze different Arctic drivers of midlatitude winter circulation," *Journal of Climate*, vol. 29, no. 11, pp. 4069–4081, 2016.

[5] M. Kretschmer, J. Runge, and D. Coumou, "Early prediction of extreme stratospheric polar vortex states based on causal precursors: Prediction of extreme vortex states," *Geophysical research letters*, vol. 44, no. 16, pp. 8592–8600, 2017.

[6] N. Jajcay, S. Kravtsov, G. Sugihara, A. A. Tsonis, and M. Palus, "Synchronization and causality across time scales in El Niño Southern Oscillation," *npj Climate and Atmospheric Science*, vol. 1, no. 1, p. 33, 2018.

[7] G. Di Capua, M. Kretschmer, J. Runge, A. Alessandri, R. V. Donner, B. van den Hurk, R. Vellore, R. Krishnan, and D. Coumou, "Long-lead statistical forecasts of the indian summer monsoon rainfall based on causal precursors," *Weather and Forecasting*, vol. 34, no. 5, pp. 1377–1394, 2019.

[8] E. A. Barnes, S. M. Samarasinghe, I. Ebert-Uphoff, and J. C. Furtado, "Tropospheric and stratospheric causal pathways between the MJO and NAO," *Journal of Geophysical Research: Atmospheres*, vol. 124, no. 16, pp. 9356–9371, 2019.

[9] J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M. D. Mahecha, J. Muñoz-Marí, E. H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Schölkopf, P. Spirtes, G. Sugihara, J. Sun, K. Zhang, and J. Zscheischler, "Inferring causation from time series in Earth system sciences.," *Nature Communications*, vol. 10, no. 1, p. 2553, 2019.

[10] M. C. McGraw and E. A. Barnes, "New insights on subseasonal Arctic–midlatitude causal connections from a regularized regression model," *Journal of Climate*, vol. 33, no. 1, pp. 213–228, 2020.

[11] A. Vázquez-Patiño, L. Campozano, D. Mendoza, and E. Samaniego, "A causal flow approach for the evaluation of global climate models," *International Journal of Climatology*, 2020.

[12] A. H. Baker, D. M. Hammerling, S. A. Mickelson, H. Xu, M. B. Stolpe, P. Naveau, B. Sanderson, I. Ebert-Uphoff, S. Samarasinghe, F. De Simone, F. Carbone, C. N. Gencarelli, J. M. Dennis, J. E. Kay, and P. Lindstrom, "Evaluating lossy data compression on climate simulation data within a large ensemble," *Geoscientific Model Development*, vol. 9, no. 12, pp. 4381–4403, 2016.

[13] P. Nowack and J. Runge, "Large-scale causal network discovery in CMIP5 models: robustness and intercomparison," in *AGU Fall Meeting Abstracts*, 2018. `https://agu.confex.com/agu/fm18/meetingapp.cgi/Paper/411521`.

[14] Y. Deng and I. Ebert-Uphoff, "Weakening of atmospheric information flow in a warming climate in the community climate system model," *Geophysical Research Letters*, vol. 41, no. 1, pp. 193–200, 2013.

[15] A. Hannart, J. Pearl, F. Otto, P. Naveau, and M. Ghil, "Causal counterfactual theory for the attribution of weather and climate-related events," *Bulletin of the American Meteorological Society*, vol. 97, no. 1, pp. 99–110, 2016.

[16] A. Hannart, C. Vera, B. Cerne, and F. E. L. Otto, "Causal influence of anthropogenic forcings on the Argentinian heat wave of December 2013," *Bulletin of the American Meteorological Society*, vol. 96, no. 12, pp. S41–S45, 2015.

[17] M. C. McGraw, *Approaching Arctic-Midlatitude Dynamics from a Two-Way Feedback Perspective*. PhD thesis, Colorado State University, 2019.

[18] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[19] P. Spirtes, C. Meek, and T. Richardson, "Causal inference in the presence of latent variables and selection bias," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, UAI'95, (San Francisco, CA, USA), pp. 499–506, Morgan Kaufmann Publishers Inc., 1995.

[20] P. Spirtes, C. Glymour, and d. Scheines, *Causation, Prediction, and Search*. MIT Press, 2nd ed., 2000.

[21] D. Entner and P. O. Hoyer, "On causal discovery from time series data using FCI," *Proceedings of the 5th European Workshop on Probabilistic Graphical Models, PGM*, 2010.

[22] J. M. Ogarrio, P. Spirtes, and J. Ramsey, "A hybrid causal search algorithm for latent variable models," in *Proceedings of the Eighth International Conference on Probabilistic Graphical Models*, vol. 52 of *Proceedings of Machine Learning Research*, (Lugano, Switzerland), pp. 368–379, PMLR, 06–09 Sep 2016.

[23] D. Colombo, M. H. Maathuis, M. Kalisch, and T. S. Richardson, "Learning high-dimensional directed acyclic graphs with latent and selection variables," *Ann. Statist.*, vol. 40, pp. 294–321, 02 2012.

[24] S. Samarasinghe, "PC Stable example," Jul 2019. https://datasgateway.colostate.edu/resources/218.

[25] S. Samarasinghe and M. McGraw, "Bivariate Granger causality example," Nov 2019. https://datasgateway.colostate.edu/resources/234.

[26] S. M. Samarasinghe, M. C. McGraw, E. A. Barnes, and I. Ebert-Uphoff, "A study of links between the Arctic and the midlatitude jet stream using Granger and Pearl causality," *Environmetrics*, vol. 30, no. 4, p. e2540, 2018.

[27] S. M. Samarasinghe, Y. Deng, and I. Ebert-Uphoff, "A causality-based view of the interaction between synoptic- and planetary-scale atmospheric disturbances," *Journal of the Atmospheric Sciences*, vol. 77, no. 3, pp. 925–941, 2020.

[28] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufman Publishers, revised second printing ed., 1988.

[29] J. Pearl, *Causality - Models, Reasoning and Inference*. Cambridge University Press, reprinted with corrections ed., 2000.

[30] R. E. Neapolitan, *Learning Bayesian Networks*. Prentice Hall, 2003.

[31] J. Ramsey, M. Glymour, R. Sanchez-Romero, and C. Glymour, "A million variables and more: the fast greedy equivalence search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images," *International Journal of Data Science and Analytics*, vol. 3, pp. 121–129, Mar 2017.

[32] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using bayesian networks to analyze expression data," *Journal of Computational Biology*, vol. 7, no. 3-4, pp. 601–620, 2000. PMID: 11108481.

[33] S. S. Ahmed, S. Roy, and J. K. Kalita, "Assessing the effectiveness of causality inference methods for gene regulatory networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, pp. 1–1, 2018.

[34] C. Granger, "Testing for causality: A personal viewpoint," *Journal of Economic Dynamics and Control*, vol. 2, no. 1, pp. 329–352, 1980.

[35] J. Runge, V. Petoukhov, and J. Kurths, "Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models," *Journal of Climate*, vol. 27, no. 2, pp. 720–739, 2014.

[36] M. C. McGraw and E. A. Barnes, "Memory matters: A case for Granger causality in climate variability studies," *Journal of Climate*, vol. 31, pp. 3289–3300, 2018.

[37] P. Spirtes and C. Glymour, "An algorithm for fast recovery of sparse causal graphs," *Social science computer review*, vol. 9, no. 1, pp. 62–72, 1991.

[38] D. Colombo and M. H. Maathuis, "Order-independent constraint-based causal structure learning," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3741–3782, 2014.

[39] J. D. Ramsey and B. Andrews, "A comparison of public causal search packages on linear, Gaussian data with no latent variables," 2017. `http://arxiv.org/abs/1709.04240`.

[40] T. Chu, D. Danks, and C. Glymour, "Data driven methods for nonlinear Granger causality: Climate teleconnection mechanisms," tech. rep., Carnegie Mellon University, Department of Philosophy, 2005.

[41] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, "Escaping the curse of dimensionality in estimating multivariate transfer entropy," *Physical review letters*, vol. 108, no. 25, p. 258701, 2012.

[42] X. Chen, Y. Liu, H. Liu, and J. G. Carbonell, "Learning spatial-temporal varying graphs with applications to climate data analysis.," in *AAAI*, 2010.

[43] M. T. Bahadori and Y. Liu, "Granger causality analysis with hidden variables in climate science applications," in *Climate Informatics workshop (CI 2011)*, 2011.

[44] T. Zerenner, P. Friederichs, K. Lehnertz, and A. Hense, "A Gaussian graphical model approach to climate networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, no. 2, p. 023103, 2014.

[45] J. Runge, *Detecting and quantifying causality from time series of complex systems*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2014.

[46] H. Wang, A. Banerjee, C.-J. Hsieh, P. K. Ravikumar, and I. S. Dhillon, "Large scale distributed sparse precision estimation," in *Advances in Neural Information Processing Systems*, pp. 584–592, 2013.

[47] I. Ebert-Uphoff and Y. Deng, "Identifying physical interactions from climate data: Challenges and opportunities," *Computing in Science & Engineering*, vol. 17, no. 6, pp. 27–34, 2015.

[48] L. Barnett, A. B. Barrett, and A. K. Seth, "Granger causality and transfer entropy are equivalent for Gaussian variables," *Phys. Rev. Lett.*, vol. 103, p. 238701, Dec 2009.

[49] E. Kodra, S. Chatterjee, and A. R. Ganguly, "Exploring Granger causality between global average observed time series of carbon dioxide and temperature," *Theoretical and Applied Climatology*, vol. 104, no. 3, pp. 325–335, 2011.

[50] A. Attanasio, "Testing for linear Granger causality from natural/anthropogenic forcings to global temperature anomalies," *Theoretical and Applied Climatology*, vol. 110, no. 1, pp. 281–289, 2012.

[51] W. Wang, B. T. Anderson, R. K. Kaufmann, and R. B. Myneni, "The relation between the North Atlantic Oscillation and SSTs in the North Atlantic Basin," *Journal of Climate*, vol. 17, no. 24, pp. 4752–4759, 2004.

[52] T. J. Mosedale, D. B. Stephenson, M. Collins, and T. C. Mills, "Granger causality of coupled climate processes: Ocean feedback on the North Atlantic Oscillation," *Journal of Climate*, vol. 19, no. 7, pp. 1182–1194, 2006.

[53] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Springer, corr. 2nd printing ed., 2007.

[54] A. Arnold, Y. Liu, and N. Abe, "Temporal causal modeling with graphical Granger methods," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 66–75, ACM, 2007.

[55] C. Papagiannopoulou, D. G. Miralles, S. Decubber, M. Demuzere, N. E. C. Verhoest, W. A. Dorigo, and W. Waegeman, "A non-linear Granger-causality framework to investigate climate–vegetation dynamics," *Geoscientific Model Development*, vol. 10, no. 5, pp. 1945–1960, 2017.

[56] A. Tank, I. Covert, N. Foti, A. Shojaie, and E. Fox, "Neural Granger causality for nonlinear time series," 2018.

[57] A. Montalto, S. Stramaglia, L. Faes, G. Tessitore, R. Prevete, and D. Marinazzo, "Neural networks with non-uniform embedding and explicit validation phase to assess Granger causality," *Neural Networks*, vol. 71, pp. 159 – 171, 2015.

[58] F. V. Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs*. Springer, 2nd ed., 2007.

[59] R. Scheines, "An introduction to causal inference," in *Causality in Crisis? University of Notre Dame*, pp. 185–200, Press, 1997.

[60] "The Tetrad project graphical causal models." http://www.phil.cmu.edu/tetrad/. Accessed: 01/14/2020.

[61] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic, "Detecting and quantifying causal associations in large nonlinear time series datasets," *Science Advances*, vol. 5, no. 11, 2019.

[62] K. Murphy, "Bayes net toolbox for Matlab (BNT)," 2014. Available at https://github.com/bayesnet/bnt.

[63] "pcalg: Methods for graphical models and causal inference." https://cran.r-project.org/web/packages/pcalg/, as of Jan 2020.

[64] M. C. Serreze and R. G. Barry, *The Arctic Climate System*. Cambridge, UK: Cambridge University Press, 2005.

[65] E. A. Barnes and J. A. Screen, "The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?," *WIREs Clim. Change*, vol. 6, pp. 277–286, 2015.

[66] J. E. Overland and M. Wang, "Large-scale atmospheric circulation changes are associated with the recent loss of Arctic sea ice," *Tellus*, vol. 62, 2010.

[67] J. A. Francis and S. J. Vavrus, "Evidence linking Arctic amplification to extreme weather in mid-latitudes," *Geophys. Res. Lett.*, vol. 39, 2012.

[68] J. Liu, J. A. Curry, H. Wang, M. Song, and R. M. Horton, "Impact of declining Arctic sea ice on winter snowfall," *Proc. Natl. Acad. Sci. (USA)*, vol. 109, pp. 4074–4079, 2012.

[69] Q. Tang, X. Zhang, X. Yang, and J. A. Francis, "Cold winter extremes in northern continents linked to Arctic sea ice loss," *Environ. Res. Lett.*, vol. 8, 2013.

[70] J. A. Screen and I. Simmonds, "The central role of diminishing sea ice in recent Arctic temperature amplification," *Nature*, vol. 464, pp. 1334–1337, 2010.

[71] E. A. Barnes, "Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes," *Geophys. Res. Lett.*, vol. 40, 2013.

[72] E. A. Barnes, E. Dunn-Sigouin, G. Masato, and T. Woollings, "Exploring recent trends in Northern Hemisphere blocking," *Geophys. Res. Lett.*, vol. 41, 2014.

[73] G. Magnusdottir, C. Deser, and R. Saravanan, "The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part I: Main features and storm track characteristics of the response," *J. Climate*, vol. 17, pp. 857–876, 2004.

[74] A. H. Butler, D. W. J. Thompson, and R. Heikes, "The steady-state atmospheric circulation response to climate change-like thermal forcings in a simple general circulation model," *J. Climate*, vol. 23, pp. 3474–3496, 2010.

[75] C. Deser, R. A. Tomas, M. Alexander, and D. Lawrence, "The seasonal atmospheric response to projected Arctic sea ice loss in the late twenty-first century," *J. Clim*, vol. 23, pp. 333–351, 2010.

[76] Y. Peings and G. Magnusdottir, "Response of the wintertime Northern Hemisphere atmospheric circulation to current and projected Arctic sea ice decline: A numerical study with CAM5," *J. Climate*, vol. 27, pp. 244–264, 2014.

[77] R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, *et al.*, "The modern-era retrospective analysis for research and applications, version 2 (merra-2)," *Journal of Climate*, vol. 30, no. 14, pp. 5419–5454, 2017.

[78] C. Woods, R. Caballero, and G. Svensson, "Large-scale circulation associated with moisture intrusions into the Arctic during winter," *Geophys. Res. Lett.*, vol. 40, pp. 4717–4721, 2013.

[79] C. Liu and E. Barnes, "Extreme moisture transport into the Arctic linked to Rossby wave breaking," *J. Geophys. Res. Atmos.*, vol. 120, pp. 3774–3788, 2015.

[80] J. T. Overpeck, G. A. Meehl, S. Bony, and D. R. Easterling, "Climate data challenges in the 21st century," *Science*, vol. 331, no. 6018, pp. 700–702, 2011.

[81] S. Samarasinghe, M. McGraw, E. Barnes, and I. Ebert-Uphoff, "A study of causal links between the Arctic and the midlatitude jet-streams," *Proceedings of the 7th International Workshop on Climate Informatics*, 2017.

[82] G. Rebane and J. Pearl, "The recovery of causal poly-trees from statistical data," in *Proceedings of the Third Conference on Uncertainty in Artificial Intelligence*, UAI'87, (Arlington, Virginia, USA), p. 222–228, AUAI Press, 1987.

[83] C. Strong, G. Magnusdottir, and H. Stern, "Observed feedback between winter sea ice and the North Atlantic Oscillation," *Journal of Climate*, vol. 22, no. 22, pp. 6021–6032, 2009.

[84] N. J. Matthewman and G. Magnusdottir, "Observed interaction between Pacific sea ice and the Western Pacific Pattern on intraseasonal time scales," *Journal of Climate*, vol. 24, pp. 5031–5042, 2011.

[85] J. L. Thomas, D. W. Waugh, and A. Gnanadesikan, "Southern Hemisphere extratropical circulation: Recent trends and natural variability," *Geophysical Research Letters*, vol. 42, pp. 5508–5515, 2015.

[86] T. Woollings, E. Barnes, B. Hoskins, Y.-O. Kwon, R. W. Lee, C. Li, E. Madonna, M. McGraw, T. Parker, R. Rodrigues, C. Spensberger, and K. Williams, "Daily to decadal modulation of jet variability," *Journal of Climate*, vol. 31, no. 4, pp. 1297–1314, 2018.

[87] J. Kay, C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. Arblaster, S. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J.-F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein, "The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability," *Bull. Amer. Meteorol. Soc.*, vol. 96, pp. 1333–1349, 2015.

[88] R. Lund, H. Hurd, P. Bloomfield, and R. Smith, "Climatological time series with periodic correlation," *Journal of Climate*, vol. 8, pp. 2787–2809, 1995.

[89] J. M. Wallace and D. S. Gutzler, "Teleconnections in the geopotential height field during the Northern Hemisphere winter," *Mon. Weath. Rev.*, vol. 109, pp. 784–812, 1981.

[90] B. Hoskins and K. Hodges, "New perspectives on the Northern Hemisphere winter storm tracks," *J. Atmos. Sci.*, vol. 59, pp. 1041–1061, 2002.

[91] J. Wettstein and J. Wallace, "Observed patterns of month-to-month storm-track variability and their relationship to the background flow," *J. Atmos. Sci.*, vol. 67, pp. 1420–1437, 2010.

[92] T. Woollings, A. Hannachi, and B. Hoskins, "Variability of the North Atlantic eddy-driven jet stream," *Quart. J. Roy. Meteorol. Soc.*, vol. 136, pp. 856–868, 2010.

[93] E. Barnes and A. Fiore, "Surface ozone variability and the jet position: Implications for projecting future air quality," *Geophys. Res. Lett.*, vol. 40, pp. 2839–2844, 2013.

[94] B. Pfaff, "VAR, SVAR and SVEC models: implementation within R package vars," *Journal of Statistical Software*, vol. 27, no. 4, pp. 1–32, 2008.

[95] W. B. Nicholson, D. S. Matteson, and J. Bien, "VARX-L: Structured regularization for large vector autoregressions with exogenous variables," *International Journal of Forecasting*, vol. 33, no. 3, pp. 627–651, 2017.

[96] V. Ivanov and L. Kilian, "A practitioner's guide to lag order selection for VAR impulse response analysis," *Studies in Nonlinear Dynamics & Econometrics*, vol. 9, no. 1, 2005.

[97] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[98] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman and Hall/CRC, 2015.

[99] J. Qian, T. Hastie, J. Friedman, R. Tibshirani, and N. Simon, "Glmnet for Matlab," 2013. Available at `http://www.stanford.edu/~hastie/glmnet_matlab/`.

[100] L. Melkumova and S. Shatskikh, "Comparing ridge and lasso estimators for data analysis," *Procedia Engineering*, vol. 201, pp. 746–755, 2017.

[101] E. A. Barnes and D. L. Hartmann, "Detection of Rossby wave breaking and its response to shifts of the midlatitude jet with climate change," *J. Geophys. Res.: Atmos.*, vol. 117, 2012.

[102] J. Runge, "Causal network reconstruction from time series: From theoretical assumptions to practical estimation," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 7, p. 075310, 2018.

[103] D. W. J. Thompson and J. M. Wallace, "The Arctic oscillation signature in the wintertime geopotential height and temperature fields," *Geophysical Research Letters*, vol. 25, no. 9, pp. 1297–1300, 1998.

[104] D. Thompson and J. Wallace, "Annular modes in the extratropical circulation. Part I: Month-to-month variability," *J. Climate*, vol. 13, pp. 1000–1016, 2000.

[105] G. T. Walker and E. W. Bliss, "World weather v," *Memoirs of the Royal Meteorological Society*, vol. 4, pp. 53–84, 1932.

[106] H. van Loon and J. C. Rogers, "The Seesaw in Winter Temperatures between Greenland and Northern Europe. Part I: General Description," *Monthly Weather Review*, vol. 106, no. 3, pp. 296–310, 1978.

[107] J. W. Hurrell, "Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperature," *Geophysical Research Letters*, vol. 23, no. 6, pp. 665–668, 1996.

[108] D. Thompson and J. Wallace, "Regional climate impacts of the Northern Hemisphere annular mode," *Science*, vol. 293, pp. 85–89, 2001.

[109] F. Vitart, A. W. Roberston, and D. L. T. Anderson, "Subseasonal to Seasonal Prediction Project: Bridging the gap between weather and climate," *WMO Bull.*, vol. 61, pp. 23–28, 2012.

[110] B. Hoskins, "The potential for skill across the range of the seamless weather–climate prediction problem: A stimulus for our science," *Quart. J. Roy. Meteor. Soc.*, vol. 139, pp. 573–584, 2013.

[111] A. A. Scaife, A. Arribas, E. Blockley, A. Brookshaw, R. T. Clark, N. Dunstone, R. Eade, D. Fereday, C. K. Folland, M. Gordon, L. Hermanson, J. R. Knight, D. J. Lea, C. MacLachlan, A. Maidens, M. Martin, A. K. Peterson, D. Smith, M. Vellinga, E. Wallace, J. Waters, and A. Williams, "Skillful long-range prediction of European and North American winters," *Geophysical Research Letters*, vol. 41, no. 7, pp. 2514–2519, 2014.

[112] A. Eliassen, "Slow thermally or frictionally controlled meridional circulation in a circular vortex," *Astrophysica Norvegica*, vol. 5, pp. 19–60, 1950.

[113] P. Haynes, C. Marks, M. McIntyre, T. Shepherd, and K. Shine, "On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces," *J. Atmos. Sci.*, vol. 48, pp. 651–678, 1991.

[114] M. P. Baldwin, X. Cheng, and T. J. Dunkerton, "Observed correlations between winter-mean tropospheric and stratospheric circulation anomalies," *Geophysical Research Letters*, vol. 21, no. 12, pp. 1141–1144, 1994.

[115] M. P. Baldwin and T. J. Dunkerton, "Stratospheric Harbingers of Anomalous Weather Regimes," *Science*, vol. 294, no. 5542, 2001.

[116] J. Kidston, A. A. Scaife, S. C. Hardiman, D. M. Mitchell, N. Butchart, M. P. Baldwin, and L. J. Gray, "Stratospheric influence on tropospheric jet streams, storm tracks and surface weather," *Nature Publishing Group*, vol. 8, pp. 433–440, 2015.

[117] Chen. P. and W. A. Robinson, "Propagation of planetary waves between the troposphere and stratosphere," *J. Atmos. Sci.*, vol. 49, pp. 2533–2544, 1992.

[118] P. J. Kushner and L. M. Polvani, "Stratosphere—Troposphere Coupling in a Relatively Simple AGCM: The Role of Eddies," *Journal of Climate*, vol. 17, no. 3, pp. 629–639, 2004.

[119] C. Garfinkel, D. Waugh, and E. Gerber, "The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere," *J. Climate*, vol. 26, pp. 2077–2095, 2013.

[120] A. Scaife, A. Yu Karpechko, M. Baldwin, A. Brookshaw, A. Butler, R. Eade, M. Gordon, C. MacLachlan, N. Martin, N. Dunstone, and D. Smith, "Seasonal winter forecasts and the stratosphere," *Atmos. Sci. Let*, vol. 17, pp. 51–56, 2016.

[121] A. W. Robertson, A. Kumar, M. Peña, and F. Vitart, "Improving and Promoting Subseasonal To Seasonal Prediction," *Bulletin of the American Meteorological Society*, vol. 96, pp. ES49–ES53, 2015.

[122] R. Madden and P. Julian, "Description of a 40-50 day oscillation in the zonal wind in the tropical Pacific," *J. Atmos. Sci.*, vol. 28, pp. 702–708, 1971.

[123] B. Hoskins and D. Karoly, "The steady linear response of a spherical atmosphere to thermal and orographic forcing," *J. Atmos. Sci.*, vol. 38, pp. 1179–1196, 1981.

[124] P. D. Sardeshmukh and B. J. Hoskins, "The Generation of Global Rotational Flow by Steady Idealized Tropical Divergence," *Journal of the Atmospheric Sciences*, vol. 45, no. 7, pp. 1228–1251, 1988.

[125] J. S. Frederiksen and H. Lin, "Tropical—Extratropical Interactions of Intraseasonal Oscillations," *Journal of the Atmospheric Sciences*, vol. 70, no. 10, pp. 3180–3197, 2013.

[126] T. R. Knutson and K. M. Weickmann, "30—60 Day Atmospheric Oscillations: Composite Life Cycles of Convection and Circulation Anomalies," *Monthly Weather Review*, vol. 115, no. 7, pp. 1407–1436, 1987.

[127] L. Ferranti, T. N. Palmer, F. Molteni, and E. Klinker, "Tropical-Extratropical Interaction Associated with the 30-60 Day Oscillation and Its Impact on Medium and Extended Range Prediction," *Journal of the Atmospheric Sciences*, vol. 47, no. 18, pp. 2177–2199, 1990.

[128] R. W. Moore, O. Martius, and T. Spengler, "The Modulation of the Subtropical and Extratropical Atmosphere in the Pacific Basin in Response to the Madden-Julian Oscillation," *Monthly Weather Review*, vol. 138, pp. 2761–2779, 2010.

[129] M. L'Heureux and W. Higgins, "Boreal winter links between the Madden-Julian Oscillation and the Arctic Oscillation," *J. Climate*, vol. 21, pp. 3040–3050, 2008.

[130] H. Lin, G. Brunet, and J. Derome, "An Observed Connection between the North Atlantic Oscillation and the Madden—Julian Oscillation," *Journal of Climate*, vol. 22, no. 2, pp. 364–380, 2009.

[131] Z. Jiang, S. B. Feldstein, and S. Lee, "The relationship between the Madden-Julian Oscillation and the North Atlantic Oscillation," *Quarterly Journal of the Royal Meteorological Society*, vol. 143, no. 702, pp. 240–250, 2017.

[132] P. Yadav and D. M. Straus, "Circulation Response to Fast and Slow MJO Episodes," *Monthly Weather Review*, vol. 145, no. 5, pp. 1577–1596, 2017.

[133] C. Cassou, "Intraseasonal interaction between the Madden—Julian Oscillation and the North Atlantic Oscillation," *Nature*, vol. 455, pp. 523–527, 2008.

[134] H. Lin, G. Brunet, and J. S. Fontecilla, "Impact of the Madden-Julian Oscillation on the intraseasonal forecast skill of the North Atlantic Oscillation," *Geophysical Research Letters*, vol. 37, no. 19, pp. n/a–n/a, 2010.

[135] P. E. Roundy, K. MacRitchie, J. Asuma, and T. Melino, "Modulation of the Global Atmospheric Circulation by Combined Activity in the Madden—Julian Oscillation and the El Niño—Southern Oscillation during Boreal Winter," *Journal of Climate*, vol. 23, no. 15, pp. 4045–4059, 2010.

[136] B. Pang, Z. Chen, Z. Wen, and R. Lu, "Impacts of two types of El Niño on the MJO during boreal winter," *Advances in Atmospheric Sciences*, vol. 33, no. 8, pp. 979–986, 2016.

[137] S.-W. Son, Y. Lim, C. Yoo, H. H. Hendon, J. Kim, S.-W. Son, Y. Lim, C. Yoo, H. H. Hendon, and J. Kim, "Stratospheric Control of the Madden—Julian Oscillation," *Journal of Climate*, vol. 30, no. 6, pp. 1909–1922, 2017.

[138] B. C. Weare, "Tropospheric-stratospheric wave propagation during El Niño-Southern Oscillation," *Journal of Geophysical Research*, vol. 115, no. D18, p. D18122, 2010.

[139] C. I. Garfinkel, S. B. Feldstein, D. W. Waugh, C. Yoo, and S. Lee, "Observed connection between stratospheric sudden warmings and the Madden-Julian Oscillation," *Geophysical Research Letters*, vol. 39, 2012.

[140] C. I. Garfinkel, J. J. Benedict, and E. D. Maloney, "Impact of the MJO on the boreal winter extratropical circulation," *Geophysical Research Letters*, vol. 41, no. 16, pp. 6055–6062, 2014.

[141] C. Liu, B. Tian, K.-F. Li, G. L. Manney, N. J. Livesey, Y. L. Yung, and D. E. Waliser, "Northern Hemisphere mid-winter vortex-displacement and vortex-split stratospheric sudden warmings: Influence of the Madden-Julian Oscillation and Quasi-Biennial Oscillation," *Journal of Geophysical Research: Atmospheres*, vol. 119, pp. 12,599–12,620, 2014.

[142] W. Kang and E. Tziperman, "More Frequent Sudden Stratospheric Warming Events due to Enhanced MJO Forcing Expected in a Warmer Climate," *Journal of Climate*, vol. 30, no. 21, pp. 8727–8743, 2017.

[143] C. I. Garfinkel and C. Schwartz, "MJO-Related Tropical Convection Anomalies Lead to More Accurate Stratospheric Vortex Variability in Subseasonal Forecast Models," *Geophysical Research Letters*, vol. 44, no. 19, pp. 10,054–10,062, 2017.

[144] C. Schwartz and C. I. Garfinkel, "Relative roles of the MJO and stratospheric variability in North Atlantic and European winter climate," *Journal of Geophysical Research: Atmospheres*, vol. 122, pp. 4184–4201, 2017.

[145] M. C. Wheeler and H. H. Hendon, "An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction," *Monthly Weather Review*, vol. 132, no. 8, pp. 1917–1932, 2004.

[146] D. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. Geer, L. Haimberger, S. Healy, H. Hersbach, E. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. McNally, B. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, "The ERA-Interim reanalysis: configuration and performance of the data assimilation system," *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 656, pp. 553–597, 2011.

[147] A. J. Charlton-Perez, L. Ferranti, and R. W. Lee, "The influence of the stratospheric state on North Atlantic weather regimes," *Quarterly Journal of the Royal Meteorological Society*, vol. 104, p. 30937, Aug. 2018.

[148] V. Limpasuvan, D. W. J. Thompson, and D. L. Hartmann, "The Life Cycle of the Northern Hemisphere Sudden Stratospheric Warmings," *Journal of Climate*, vol. 17, pp. 2584–2596, 2004.

[149] W. J. M. Seviour, "Weakening and shift of the Arctic stratospheric polar vortex: Internal variability or forced response?," *Geophysical Research Letters*, vol. 44, pp. 3365–3373, apr 2017.

[150] CPC, "Monitoring & Data: Daily North Atlantic Oscillation Index." `https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao_index.html`, 2018.

[151] A. G. Barnston and R. E. Livezey, "Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns," *Monthly Weather Review*, vol. 115, no. 6, pp. 1083–1126, 1987.

[152] C. Stan, D. M. Straus, J. S. Frederiksen, H. Lin, E. D. Maloney, and C. Schumacher, "Review of Tropical-Extratropical Teleconnections on Intraseasonal Time Scales," *Reviews of Geophysics*, vol. 55, pp. 902–937, 2017.

[153] T. Chu and C. Glymour, "Search for additive nonlinear time series causal models," *Journal of Machine Learning Research*, vol. 9, pp. 967–991, 2009.

[154] D. Malinsky and D. Danks, "Causal discovery algorithms: A practical guide," *Philosophy Compass*, vol. 13, no. 1, p. e12470, 2018.

[155] D. W. J. Thompson, M. P. Baldwin, and J. M. Wallace, "Stratospheric Connection to Northern Hemisphere Wintertime Weather: Implications for Prediction," *J. Climate*, vol. 15, pp. 1421–1428, 2002.

[156] M. H. P. Ambaum and B. J. Hoskins, "The NAO Troposphere-Stratosphere Connection," *Journal of Climate*, vol. 15, pp. 1969–1978, 2002.

[157] A. Butler, A. Charlton-Perez, D. I. Domeisen, C. Garfinkel, E. P. Gerber, P. Hitchcock, A. Y. Karpechko, A. C. Maycock, M. Sigmond, I. Simpson, and S.-W. Son, "Sub-seasonal Predictability and the Stratosphere," in *Sub-Seasonal to Seasonal Prediction*, pp. 223–241, Elsevier, 2019.

[158] D. M. Mitchell, L. J. Gray, J. Anstey, M. P. Baldwin, and A. J. Charlton-Perez, "The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate," *Journal of climate*, vol. 26, pp. 2668–2682, Apr. 2013.

[159] S. Zhou and A. J. Miller, "The Interaction of the Madden—Julian Oscillation and the Arctic Oscillation," *Journal of Climate*, vol. 18, pp. 143 – 159, 2005.

[160] S. Zhou, M. L'Heureux, S. Weaver, and A. Kumar, "A composite study of the MJO influence on the surface air temperature and precipitation over the Continental United States," *Climate Dynamics*, vol. 38, pp. 1459–1471, 2012.

[161] K. Kodera, "Influence of stratospheric sudden warming on the equatorial troposphere," *Geophysical Research Letters*, vol. 33, no. 6, p. L06804, 2006.

[162] M. Gómez-Escolar, N. Calvo, D. Barriopedro, and S. Fueglistaler, "Tropical response to stratospheric sudden warmings and its modulation by the QBO," *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 12, pp. 7382–7395, 2014.

[163] B. Pohl and A. J. Matthews, "Observed Changes in the Lifetime and Amplitude of the Madden-Julian Oscillation Associated with Interannual ENSO Sea Surface Temperature Anomalies," *Journal of Climate*, vol. 20, no. 11, pp. 2659–2674, 2007.

[164] A. G. Marshall, H. H. Hendon, and G. Wang, "On the role of anomalous ocean surface temperatures for promoting the record Madden-Julian Oscillation in March 2015," *Geophysical Research Letters*, vol. 43, no. 1, pp. 472–481, 2016.

[165] C. Yoo and S.-W. Son, "Modulation of the boreal wintertime Madden-Julian oscillation by the stratospheric quasi-biennial oscillation," *Geophysical Research Letters*, vol. 43, no. 3, pp. 1392–1398, 2016.

[166] M. Kretschmer, J. Cohen, V. Matthias, J. Runge, and D. Coumou, "The different stratospheric influence on cold-extremes in Eurasia and North America," *npj Climate and Atmospheric Science*, vol. 1, no. 1, p. 44, 2018.

[167] J. S. A. Green, "The weather during july 1976: Some dynamical considerations of the drought," *Weather*, vol. 32, no. 4, pp. 120–126, 1977.

[168] H. Nakamura and J. Wallace, "Synoptic behavior of baroclinic eddies during the blocking onset," *Monthly Weather Review - MON WEATHER REV*, vol. 121, 07 1993.

[169] S. Maeda, C. Kobayashi, K. Takano, and T. Tsuyuki, "Relationship between singular modes of blocking flow and high-frequency eddies," *Journal of the Meteorological Society of Japan*, vol. 78, pp. 631–646, 10 2000.

[170] T.-W. Park, Y. Deng, W. Li, S. Yang, and M. Cai, "Mass footprints of the north pacific atmospheric blocking highs," *Journal of Climate*, vol. 28, no. 12, pp. 4941–4949, 2015.

[171] M. Jiwang and X. S. Liang, "Multiscale dynamical processes underlying the wintertime atlantic blockings," *Journal of the Atmospheric Sciences*, vol. 74, 07 2017.

[172] R. M. Dole and R. X. Black, "Life cycles of persistent anomalies. part ii: The development of persistent negative height anomalies over the north pacific ocean," *Monthly Weather Review*, vol. 118, no. 4, pp. 824–846, 1990.

[173] N.-C. Lau and E. O. Holopainen, "Transient eddy forcing of the time-mean flow as identified by geopotential tendencies," *Journal of the Atmospheric Sciences*, vol. 41, no. 3, pp. 313–328, 1984.

[174] M. Cai and M. Mak, "Symbiotic relation between planetary and synoptic-scale waves," *Journal of the Atmospheric Sciences*, vol. 47, no. 24, pp. 2953–2968, 1990.

[175] M. Cai and H. M. Van Den Dool, "Low-frequency waves and traveling storm tracks. part I: Barotropic component," *Journal of the Atmospheric Sciences*, vol. 48, no. 11, pp. 1420–1436, 1991.

[176] M. Cai and H. M. Van Den Dool, "Frequency waves and traveling storm tracks. part II: Three-dimensional structure," *Journal of the Atmospheric Sciences*, vol. 49, no. 24, pp. 2506–2524, 1992.

[177] W. A. Robinson, "The dynamics of the zonal index in a simple model of the atmosphere," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 43, no. 5, pp. 295–305, 1991.

[178] W. A. Robinson, "Does eddy feedback sustain variability in the zonal index?," *Journal of the Atmospheric Sciences*, vol. 53, no. 23, pp. 3556–3569, 1996.

[179] J.-Y. Yu and D. L. Hartmann, "Zonal flow vacillation and eddy forcing in a simple GCM of the atmosphere," *Journal of the Atmospheric Sciences*, vol. 50, no. 19, pp. 3244–3259, 1993.

[180] D. J. Lorenz and D. L. Hartmann, "Eddy–zonal flow feedback in the southern hemisphere," *Journal of the Atmospheric Sciences*, vol. 58, no. 21, pp. 3312–3327, 2001.

[181] Y. Hayashi, "Estimation of nonlinear energy transfer spectra by the cross-spectral method," *Journal of the Atmospheric Sciences*, vol. 37, no. 2, pp. 299–307, 1980.

[182] J. Sheng and Y. Hayashi, "Observed and simulated energy cycles in the frequency domain," *Journal of The Atmospheric Sciences - J ATMOS SCI*, vol. 47, pp. 1243–1254, 05 1990.

[183] J. Sheng and D. Jacques, "An observational study of the energy transfer between the seasonal mean flow and transient eddies," *Tellus A*, vol. 43, no. 2, pp. 128–144, 1991.

[184] J. Sheng and J. Derome, "Dynamic forcing of the slow transients by synoptic-scale eddies: An observational study," *Journal of the Atmospheric Sciences*, vol. 50, no. 5, pp. 757–771, 1993.

[185] M. Cai and H. M. Van Den Dool, "Dynamical decomposition of low-frequency tendencies," *Journal of the Atmospheric Sciences*, vol. 51, no. 14, pp. 2086–2100, 1994.

[186] T. J. Cuff and M. Cai, "Interaction between the low- and high-frequency transients in the southern hemisphere winter circulation," *Tellus A*, vol. 47, no. 3, pp. 331–350, 1995.

[187] Y. Deng and T. Jiang, "Intraseasonal modulation of the north pacific storm track by tropical convection in boreal winter," *Journal of Climate*, vol. 24, no. 4, pp. 1122–1137, 2011.

[188] T. Jiang, Y. Deng, and W. Li, "Local kinetic energy budget of high-frequency and intermediate-frequency eddies: winter climatology and interannual variability," *Climate Dynamics*, vol. 41, no. 3, pp. 961–976, 2013.

[189] T. Jiang, K. J. Evans, Y. Deng, and X. Dong, "Intermediate frequency atmospheric disturbances: A dynamical bridge connecting western U.S. extreme precipitation with East Asian cold surges," *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 7, pp. 3723–3735, 2014.

[190] X. San Liang, "Canonical transfer and multiscale energetics for primitive and quasi-geostrophic atmospheres," *Journal of the Atmospheric Sciences*, vol. 73, no. 11, pp. 4439–4468, 2016.

[191] F. Xu and X. San Liang, "On the generation and maintenance of the 2012/13 sudden stratospheric warming," *Journal of the Atmospheric Sciences*, vol. 74, no. 10, pp. 3209–3228, 2017.

[192] X. S. Liang and R. Kleeman, "A rigorous formalism of information transfer between dynamical system components. i. discrete mapping," *Physica D: Nonlinear Phenomena*, vol. 231, no. 1, pp. 1 – 9, 2007.

[193] X. S. Liang and R. Kleeman, "A rigorous formalism of information transfer between dynamical system components. ii. continuous flow," *Physica D: Nonlinear Phenomena*, vol. 227, no. 2, pp. 173 – 182, 2007.

[194] X. S. Liang, "Information flow and causality as rigorous notions ab initio," *Phys. Rev. E*, vol. 94, p. 052201, Nov 2016.

[195] M. Paluš, "Cross-scale interactions and information transfer," *Entropy*, vol. 16, p. 5263–5289, Oct 2014.

[196] P. Spirtes, C. Glymour, and R. Scheines. New York, NY: Springer New York, 1993.

[197] D. Koller and N. Friedman, *Probabilistic Graphical Models - Principles and Techniques*. MIT Press, 1st ed., 2009.

[198] B. Abramson, J. Brown, W. Edwards, A. Murphy, and R. L. Winkler, "Hailfinder: A bayesian system for forecasting severe weather," *International Journal of Forecasting*, vol. 12, no. 1, pp. 57 – 71, 1996. Probability Judgmental Forecasting.

[199] M. Cossentino, F. Raimondi, M. Vitale, and P. Italy *WIT Transactions on Ecology and the Environment*, vol. 47.

[200] A. S. Cofiño, R. Cano, C. Sordo, and J. M. Gutiérrez, "Bayesian networks for probabilistic weather prediction," in *ECAI*, 2002.

[201] J. Runge, J. Heitzig, N. Marwan, and J. Kurths, "Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 86, p. 061121, 12 2012.

[202] J. Hlinka, D. Hartman, M. Vejmelka, J. Runge, N. Marwan, J. Kurths, and M. Paluš, "Reliability of inference of directed climate networks using conditional mutual information," *Entropy*, vol. 15, no. 6, pp. 2023–2045, 2013.

[203] M. Kretschmer, D. Coumou, J. F. Donges, and J. Runge, "Using Causal Effect Networks to Analyze Different Arctic Drivers of Midlatitude Winter Circulation," *Journal of climate*, vol. 29, no. 11, pp. 4069–4081, 2016.

[204] Y. Deng and I. Ebert-Uphoff, "Weakening of atmospheric information flow in a warming climate in the community climate system model," *Geophysical Research Letters*, vol. 41, no. 1, pp. 193–200, 2014.

[205] I. Ebert-Uphoff and Y. Deng, "Causal discovery in the geosciences using synthetic data to learn how to interpret results," *Computers and Geosciences*, vol. 99, pp. 50 – 60, 2017.

[206] X. Liang, "The Liang-Kleeman information flow: Theory and applications," *Entropy*, vol. 15, p. 327–360, Jan 2013.

[207] M. Materassi, G. Consolini, N. Smith, and R. De Marco, "Information theory analysis of cascading process in a synthetic model of fluid turbulence," *Entropy*, vol. 16, p. 1272–1286, Feb 2014.

[208] X. S. Liang and A. Lozano-Durán, "A preliminary study of the causal structure in fully developed near-wall turbulence," *Proceedings of the Summer Program 2016, Center for Turbulence Research, Stanford University, CA*, vol. 2016, pp. 233–242, 12 2016.

[209] X. S. Liang, "A study of the cross-scale causation and information flow in a stormy model mid-latitude atmosphere," *Entropy*, vol. 21, p. 149, 02 2019.

[210] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, "The NCEP/NCAR 40-year reanalysis project," *Bulletin of the American Meteorological Society*, vol. 77, no. 3, pp. 437–472, 1996.

[211] R. Kistler, E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, "The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation," *Bulletin of the American Meteorological Society*, vol. 82, no. 2, pp. 247–268, 2001.

[212] M. L. Blackmon, "A climatological spectral study of the 500 mb geopotential height of the northern hemisphere," *Journal of the Atmospheric Sciences*, vol. 33, no. 8, pp. 1607–1623, 1976.

[213] S. Samarasinghe, Y. Deng, and I. Ebert-Uphoff, "Structure learning in spectral space with applications in climate science," in *Proc. Workshop on Mining Big Data in Climate and Environment (MBDCE 2017), 17th SIAM Int. Conf. on Data Mining (SDM 2017), Houston, TX.*, 2017.

[214] Y. Deng and M. Mak, "Nature of the differences in the intraseasonal variability of the pacific and atlantic storm tracks: A diagnostic study," *Journal of the Atmospheric Sciences*, vol. 63, no. 10, pp. 2602–2615, 2006.

[215] M. Mak and Y. Deng, "Diagnostic and dynamical analyses of two outstanding aspects of storm tracks," *Dynamics of Atmospheres and Oceans - DYNAM ATMOS OCEANS*, vol. 43, pp. 80–99, 04 2007.

[216] E. K. M. Chang, S. Lee, and K. L. Swanson, "Storm track dynamics," *Journal of Climate*, vol. 15, no. 16, pp. 2163–2183, 2002.

[217] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, 2012.

[218] M. T. Bahadori and Y. Liu, "An examination of practical Granger causality inference," *SIAM International Conference on Data Mining*, pp. 467–475, 2013.

[219] V. Schonefeld, "Spherical harmonics," 2005. `http://limbicsoft.com/volker/prosem_paper.pdf`, as of Jan 2020.

Appendices

A.1 Details of Group LASSO Approach

In general a standard LASSO approach is adequate to perform a basic Granger analysis. In contrast to the standard LASSO approach, the *group* LASSO approach [95, 98] is often useful, if the predictors show a natural group structure, as is the case for time series data. By treating a group of coefficients as a collective entity, and constraining that group (rather than the individual coefficients), this approach forces entire groups of coefficients to be zero, rather than individual coefficients. In our case the groups are based on the predictor variables (where predictor variables = $\{\{z_{m,t-i}\}_{t=p+1}^{T} \mid i = 1, \ldots, p \; and \; m = 1, \ldots, k\}$) and the coefficients of any particular predictor variable are all set to zero if that predictor does not contribute to the predictability of *any* of the response variables. The specific version used here solves the optimization problem given in Eq. (1), where $a_{.,m}^{i}$ denotes the mth column in $\mathbf{A_i}$ and $N = (T - p)$ is the effective sample size.

$$\min_{\mathbf{A},\mathbf{c}} \quad \frac{1}{2N} \sum_{t=p+1}^{T} \|\mathbf{z_t} - \mathbf{c} - \sum_{i=1}^{p} \mathbf{A_i} \mathbf{z_{t-i}}\|_2^2 + \lambda \sum_{i=1}^{p} \sum_{m=1}^{k} \|a_{.,m}^{i}\|_2 \quad \lambda \geq 0. \tag{1}$$

A.2 Parameter Selection

This appendix discusses how the parameters were chosen for the VAR and LASSO models.

A.2.1 Choosing Maximal Lag, p

There is a trade off in selecting the maximal number of lags, p, in the VAR model. While p should be large enough to capture all the physically meaningful relationships between the variables, it should not be too large to ensure the generalization and interpretability of the model. In our study, we used two statistical model selection criteria, namely the AIC and the BIC, to identify suitable p values. The AIC and BIC values are calculated for VAR models with different values of p, using the equations 2 and 3 respectively [53, 95, 96]. Here, k denotes the number of variables

($\mathbf{z}_t \in \mathbb{R}^k$); N denotes the effective sample size and $|\tilde{\mathbf{\Sigma}}_e^i|$ denotes the determinant of the estimated error covariance matrix.

$$AIC(p) = ln|\tilde{\mathbf{\Sigma}}_e^p| + \frac{2k^2p}{N} \tag{2}$$

$$BIC(p) = ln|\tilde{\mathbf{\Sigma}}_e^p| + \frac{ln(N)k^2p}{N} \tag{3}$$

The $ln|\tilde{\mathbf{\Sigma}}_e^p|$ term in these calculations account for the accuracy of the different models while the second term ($\frac{2k^2p}{N}$ and $\frac{ln(N)k^2p}{N}$) penalizes the model based on the number of parameters to be estimated. Based on these criteria, the p values that provide low AIC/BIC values provide desirable models. The only difference between the AIC and BIC calculations is that the factor 2 in AIC is substituted with a factor $log(N)$ in BIC. When $N \geq 8$, as is the case with our results, the BIC criterion prefers simpler models with a smaller number of lags compared to AIC [96].

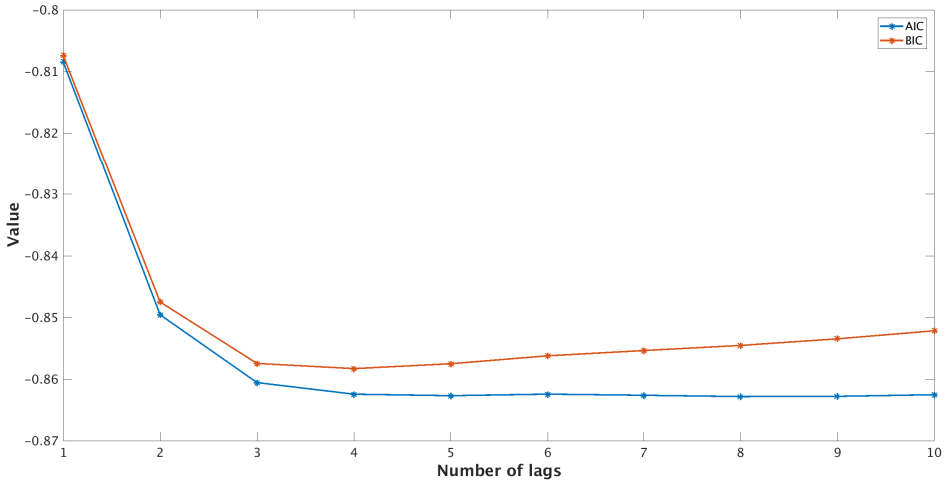
The calculated AIC and BIC values for different VAR(p) models are presented in Figure A1. Based on these values, we can see that a maximal lag between 3 and 5 is approximately suitable for both the jet speed vs. Arctic temperature model as well as the jet latitude vs. Arctic temperature model. We select a maximal lag of $p = 5$ (i.e. 25 days) for both models as it not only provides a desirable model based on the selection criteria, but also provide a time range that is interesting for our research question. We use the same number of lags for the LASSO model. Looking at higher order models with LASSO shows us that the model coefficients for the lags greater than 25 days get automatically set to zero, justifying our selection.

A.2.2 Choosing Regularization Parameter λ for LASSO

The regularization parameter λ is chosen using cross validation. Figure A2 shows the 10-fold cross validation error curves used to determine the regularization parameter, λ, in the LASSO approach. These error curves show the mean squared testing error averaged over the 10 folds along with their variability. The green dashed line corresponds to the log(λ) value associated with the minimum MSE while the blue dashed line corresponds to the error that is equal to the minimum

MSE plus one standard error. The top axis showing the degrees of freedom indicate the number of non-zero coefficients remaining in the model at specific log(λ) values.

(a)



(b)

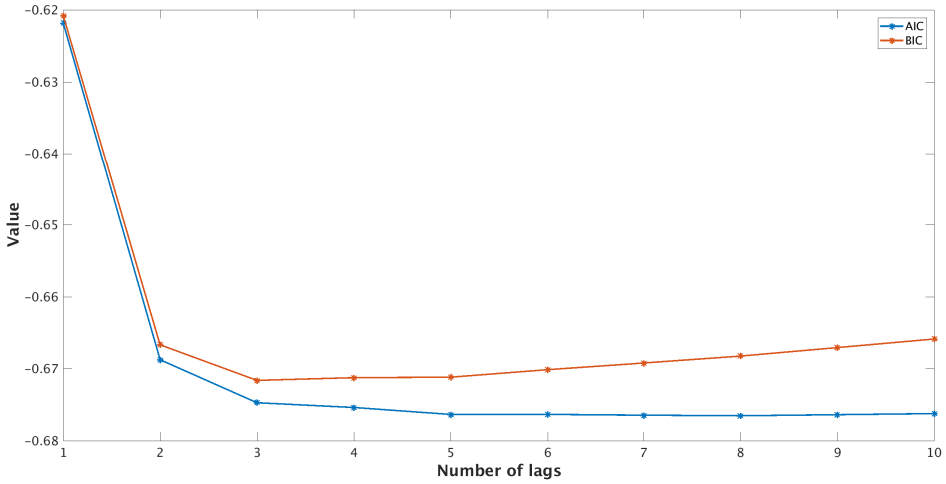


Figure A1: AIC and BIC values for VAR models between (a) Arctic temperature ($\mathcal{T}$) and jet speed ($\mathcal{S}$) and (b) Arctic temperature ($\mathcal{T}$) and jet latitude ($\mathcal{L}$). x-axis indicates the number of lags p and the y-axis indicates the AIC (blue line) and BIC (orange line) values of the VAR(p) model.

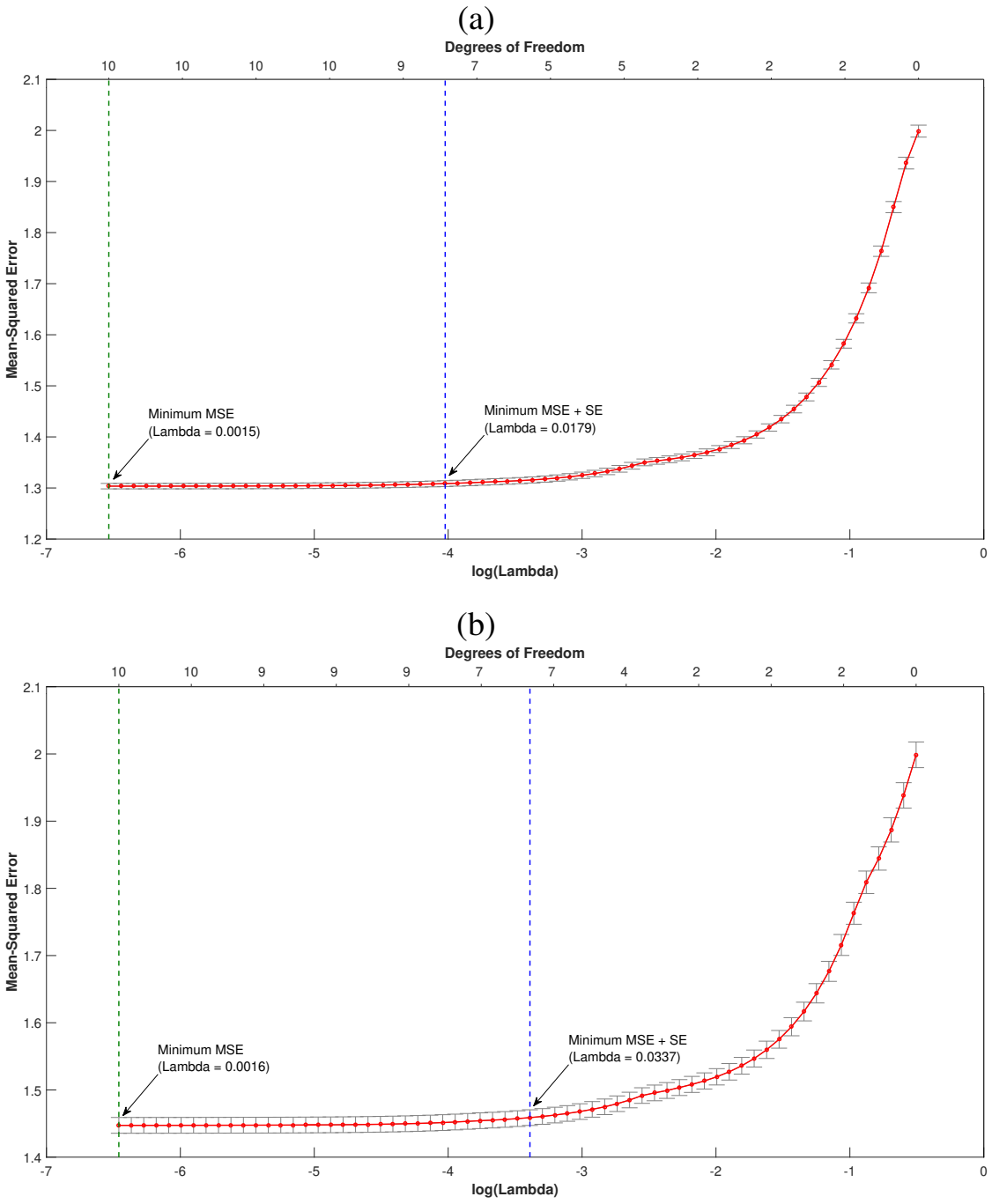


Figure A2: 10-fold cross validation error curves for LASSO for (a) Arctic temperature ($\mathcal{T}$) vs. jet speed ($\mathcal{S}$) model and (b) Arctic temperature ($\mathcal{T}$) vs. jet latitude ($\mathcal{L}$) model. x-axis indicates regularization parameter λ in logarithmic scale (bottom) and resulting number of degrees of freedom, i.e. number of non-zero regression coefficients (top). y-axis indicates the resulting mean square error from the 10-fold cross validation as average (red line) and variation (vertical intervals). Green (blue) dashed lines indicate lambda values corresponding to minimal (mimimal + one SE) error.

B.1 Parameter Details for Causal Discovery Approach

Following the notation in [1], the number of lagged copies is 13 for each variable, and we discard 1 time slice to deal with initialization issues, so we could theoretically study relationships that are up to $12 \times 5 = 60$ days apart from cause to effect. However, we only show relationships up to 35 days here. We only plot edges that have consistent connections throughout the temporal model (defined as exhibiting significant connections for at least 60% of the possible connections; see code for details). We do not allow instantaneous connections between variables, e.g., MJO(t) cannot interact with NAO(t). We use a significance value of $\alpha = 0.1$, and also examine results using $\alpha = 0.05$, for the conditional independence tests.

B.2 Supplemental Figures Showing Sample Size

The figures below are Figures 4.3–4.6 of Chapter 4 but showing sample sizes.

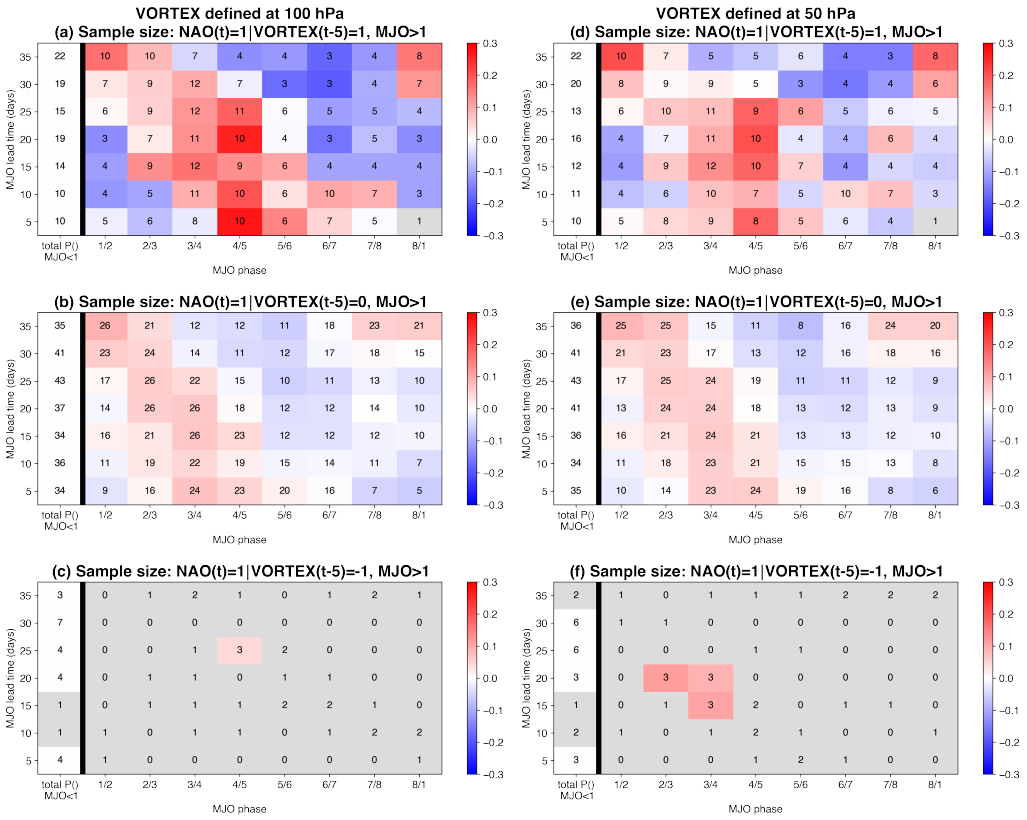


Figure B1: As in Figure 4.3 but showing sample sizes.

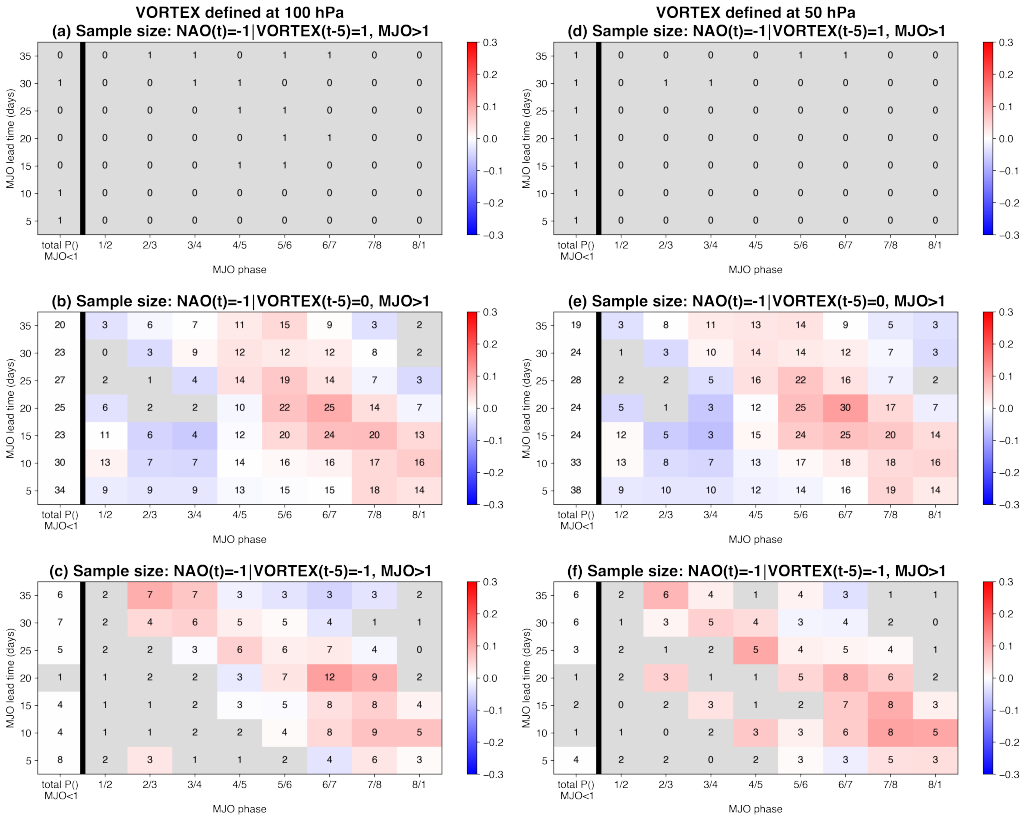


Figure B2: As in Figure 4.4 but showing sample sizes.

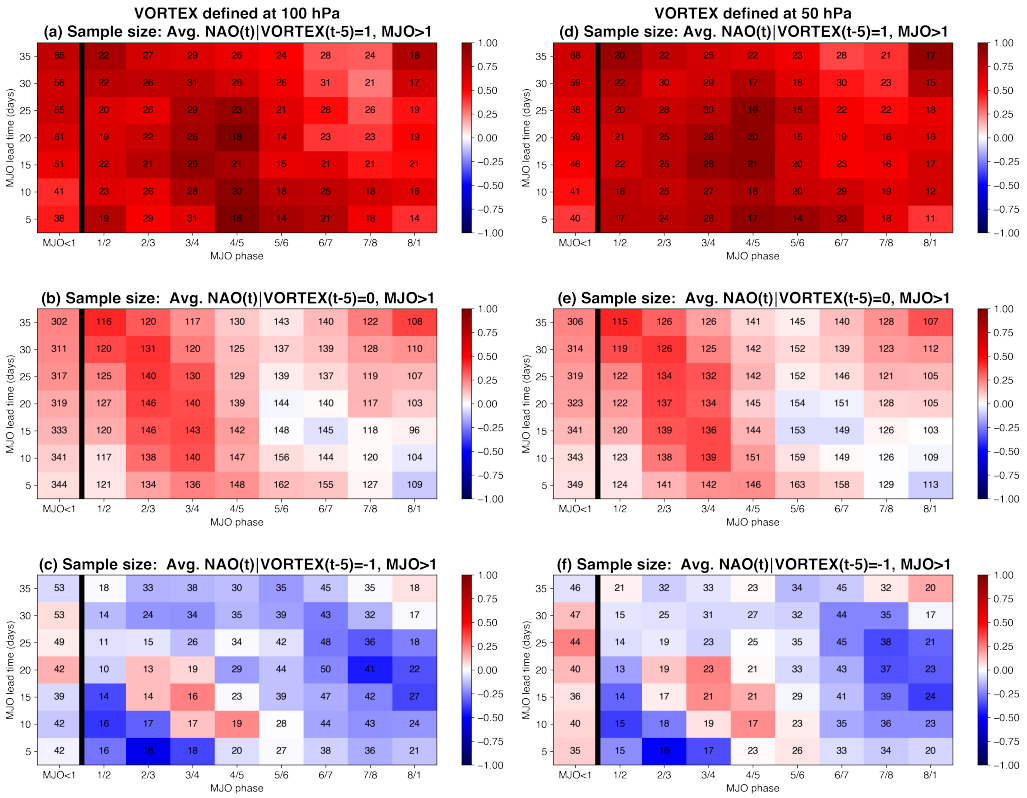


Figure B3: As in Figure 4.5 but showing sample sizes.

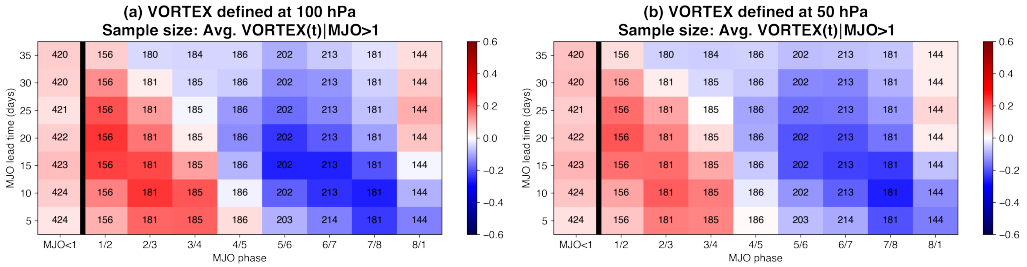


Figure B4: As in Figure 4.6 but showing sample sizes.

C.1 Review of Spherical Harmonics

Spherical harmonics provide a complete set of orthonormal basis functions that can be used to model functions on a sphere. They are defined as a combination of associated Legendre polynomials and sinusoidal functions, as follows

$$Y_L^M = N_L^M \, e^{iM\phi} \, P_L^M(cos\ \theta), \tag{4}$$

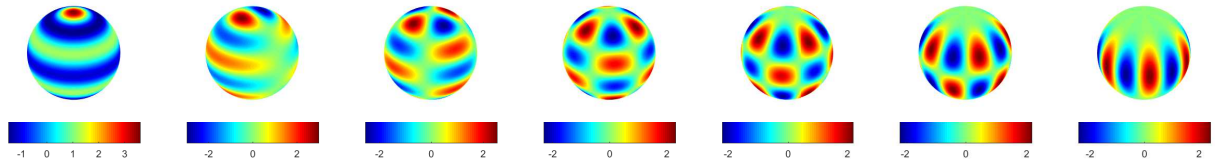
where M and L are integers, N_L^M is a normalization factor,

$$N_L^M = \left(\frac{(L-M)!(2L+1)}{4\pi(L+M)!} \right)^{1/2}, \tag{5}$$

and $P_L^M(x)$ is the associated Legendre polynomial.

$$P_L^M(x) = \frac{(-1)^M}{2^L L!} \sqrt{(1-x^2)^M} \frac{d^{(L+M)}}{dx^{(L+M)}} (x^2-1)^L. \tag{6}$$

$0 \leq \phi \leq 2\pi$ denotes the longitude angle and $0 \leq \theta \leq \pi$ denotes the polar angle, i.e. the latitude angle, but measured starting at the North pole. Fig. C1 shows basis functions for $L = 6$.



(a) R(6,0) (b) R(6,1) (c) R(6,2) (d) R(6,3) (e) R(6,4) (f) R(6,5) (g) R(6,6)

Figure C1: Real part of spherical harmonics basis functions, $R(L, M)$, for $L = 6$ and $M = 0, \ldots, 6$. [Credit: Image generated using the Matlab package *Surface Spherical Harmonic Functions Visualization* by Dimitrios Piretzidis.]

Integers L and M describe the properties of the spherical harmonics, where $L \geq 0$ and $M \in [-L, L]$. M, the *zonal wave number*, denotes the number of zero crossings in the zonal (east-west) direction, so for example there are none in Fig. C1(a). $(L - |M|)$ is the number of zero crossings

in the meridional (north-south) direction, so for example there are none in Fig. C1(g). L is called the *total wave number* and denotes the total number of zero crossings of Y_L^M on the sphere.

For any function $f(t)$ on the sphere, the spherical harmonic coefficient with indices (L, M) can be calculated by taking the inner product of integration of Y_L^M and $f(t)$ over the sphere, S:

$$C_L^M(t) =< Y_L^M, f(t) >= \int_S Y_L^M(\omega) \ \overline{f(t,\omega)} \ d\omega,$$

where the bar on top indicates the complex conjugate and ω parametrizes the integration over the sphere [219].

C.2 The Lorenz System as a Test Case for the PC Stable Algorithm

Here we present a few sample causal inference results for the Lorenz system, to showcase the scope of utility and limitations of the PC stable algorithm when used to identify cause-effect relationships. The Lorenz system, which is a simplified mathematical model for atmospheric convection, is defined by the three ordinary differential equations given below.

$$\frac{dx}{dt} = \sigma(y - x) \tag{7}$$

$$\frac{dy}{dt} = x(\rho - z) - y \tag{8}$$

$$\frac{dz}{dt} = xy - \beta z \tag{9}$$

Based on the Lorenz equations, the causality of the system involves both linear (achieved via linear terms on the right-hand side of the equation) and nonlinear (achieved via nonlinear terms on the right-hand side of the equation) causal effects: Linear effects:

1. $X(t - \delta t) \to X(t)$
2. $Y(t - \delta t) \to Y(t)$
3. $Z(t - \delta t) \to Z(t)$
4. $X(t - \delta t) \to Y(t)$
5. $Y(t - \delta t) \to X(t)$

Nonlinear effects:

1. $Z(t-\delta t) \rightarrow Y(t)$ 2. $Y(t-\delta t) \rightarrow Z(t)$ 3. $X(t-\delta t) \rightarrow Z(t)$

We generate multiple sets of realizations of the Lorenz system with different combinations of parameter values (i.e., ρ,σ and β) and present two cases here to illustrate the main findings. In case 1, we adopt the following set of parameter values and initial conditions: $\rho = 28, \sigma = 10, \beta = 8/3$, $x(0) = 1, y(0) = 1, z(0) = 1$. We use a time step of $\delta t = 0.01$. As shown in Figure C2, these parameters give rise to the classic Lorenz attractor in the x-y-z space. Figure C3 shows the time series plots and histograms of X, Y and Z. To create the temporal model, we use $X(t)$, $Y(t)$, $Z(t)$, $X(t-\delta t)$, $Y(t-\delta t)$ and $Z(t-\delta t)$ as the nodes of the graphical model. We then use Fisher's Z test on partial correlation to determine conditional independencies. We repeat this process for different sample sizes $(1000, 6000, 12,000)$, different levels of statistical significance $(0.001, 0.01, 0.05)$ and different initial conditions $([x(0), y(0), z(0)] = [1, 1, 1], [6, -7, 3], [-15, -15, 3])$. The interactions identified remain unchanged for these different parameters. A summary of the interactions is provided in Figure C4.

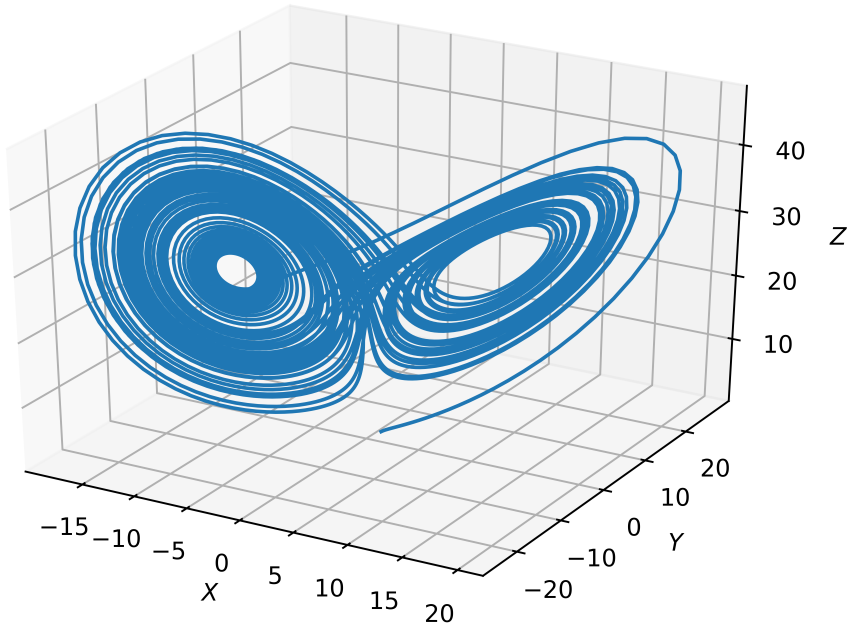


Figure C2: 3D plot of the generated data for $\rho = 28, \sigma = 10, \beta = 8/3$.

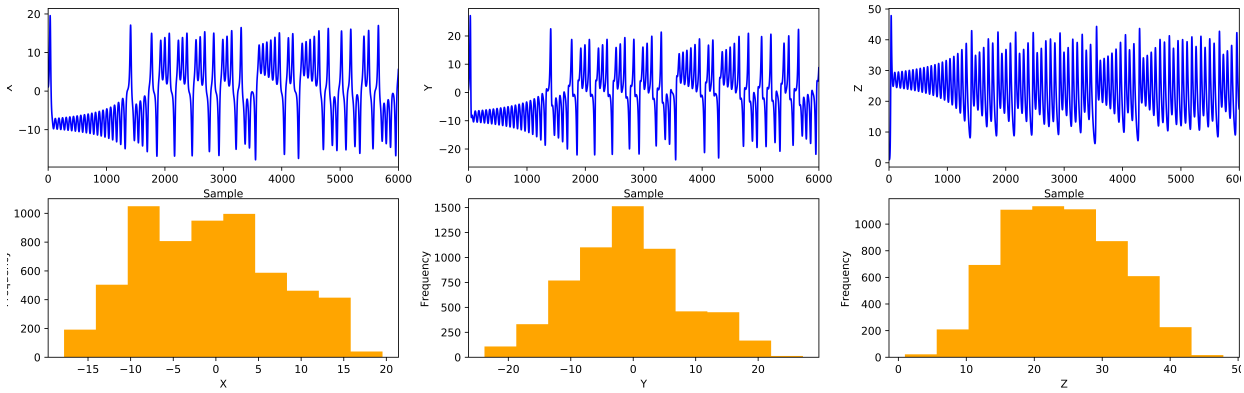


Figure C3: Time series and frequency distributions of X, Y and Z for $\rho = 28, \sigma = 10, \beta = 8/3$.

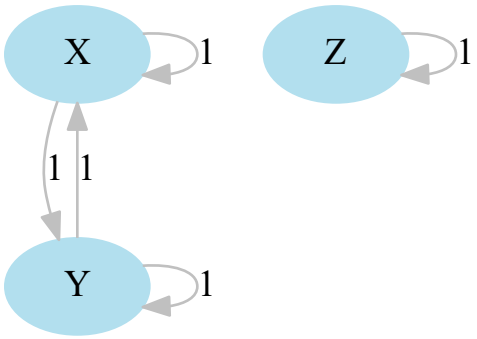


Figure C4: Summary of Interactions identified by PC stable. Here we present $1\delta t$ as 1. Parameters $\rho = 28, \sigma = 10, \beta = 8/3$.

The results indicate that the algorithm identifies all the linear causal effects as the PC stable algorithm is supposed to do. None of the nonlinear causal effects are identified. This is consistent with the fact that the conditional independence test based on partial correlation assumes that the relationships between the variables are linear. It is also important to see that the PC stable algorithm does not mistakenly identify a $Z \rightarrow X$ edge, which is an indirect connection. Note that 24.74 is a bifurcation point that changes the behavior of the Lorenz system from having a single fixed attraction point for $\rho < 24.74$ (i.e., non-chaotic) to the strange attractor with orbits around two distinct center points for $\rho > 24.74$. Based upon this, in the second case study, we look at a smaller

ρ value. This also helps to see whether the method identifies any interactions that were previously masked by a large ρ component. We generate data using the parameter values and initial conditions below: $\rho = 10, \sigma = 10, \beta = 8/3$, $x(0) = 1, y(0) = 1, z(0) = 1$. This gives rise to a stable system that converges to a fixed point with time (Fig. C5). The objective of the PC stable algorithm is to identify stochastic relationships between variables. Therefore, we only use the values of X, Y and Z at time steps prior to convergence. We use a smaller time step, δt, to ensure an adequate sample size and add to each variable a Gaussian noise with zero mean and a very small variance to ensure that all variables and their lagged variables used to create the temporal model do not hold purely deterministic relationships. Following these steps, we generate 1000 samples with a time step $\delta t = 0.005$. Figure C5 shows the trajectory of the solution while Figure C6 shows the time series plots and histograms of X, Y and Z. Figures C7 a) to c) show the summary results based on several different levels of statistical significance (α value). As we increase the value of α, the model picks up more and more expected causal relationships and finally identifies correctly all edges (both linear and nonlinear) when $\alpha = 0.2$ is used (80% level of statistical confidence). However, in contrast to the results of linear connections shown in Fig. C4, which are very robust, the results shown in Fig. C7 a) to c) are not robust. Namely, for different sample sizes and slightly different parameters, many of these edges appear or disappear.

In summary, the test cases show that the PC stable algorithm using a conditional independence test based on partial correlation can robustly identify the linear causal relationships associated with the Lorenz system while excluding indirect connections such as $Z \rightarrow X$. For a chaotic solution as in case 1, where the system spends an equal amount of time on each wing of the attractor, the changes in signs of the variables can cause some of the causal components to cancel out on average and become undetectable by an approach based on correlation/partial-correlation that measure average linear dependence. The test case shows that when the samples are taken from a regime in the (x,y,z) space where x and y do not change sign (e.g., case 2), the nonlinear dependence that was canceled out in partial correlation calculations due to changing signs of x and y (e.g., case 1) can be identified (although not robustly) when a lower level of statistical

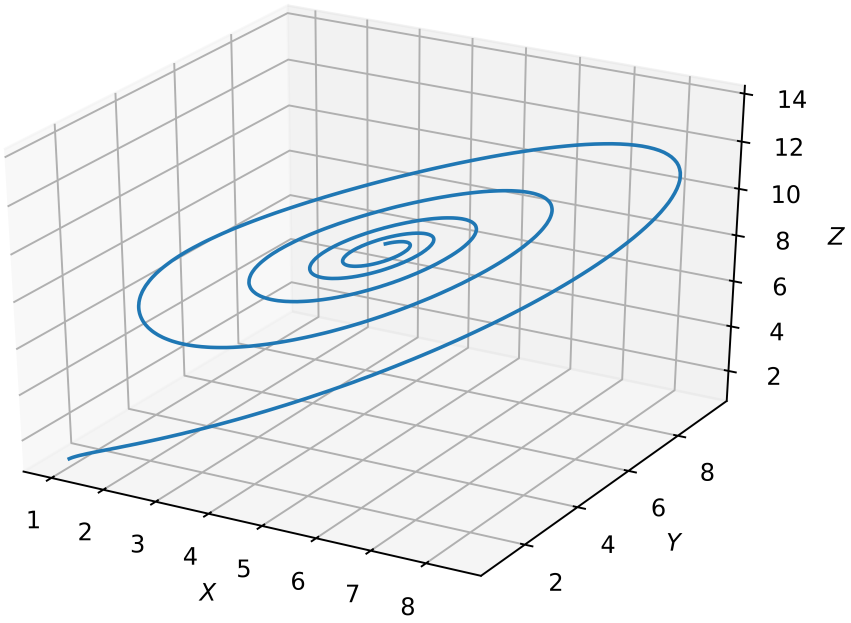


Figure C5: 3D plot of the generated data for $\rho = 10, \sigma = 10, \beta = 8/3$.

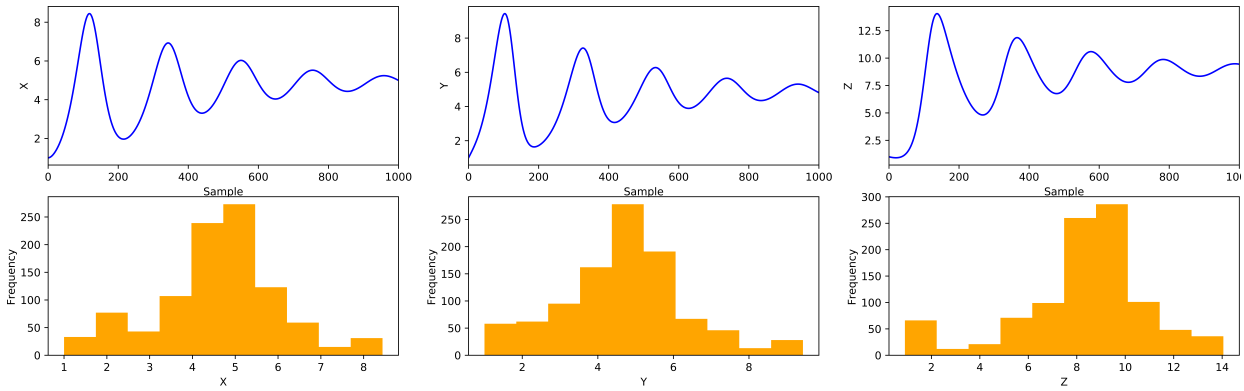


Figure C6: Time series and frequency distributions of X, Y and Z for $\rho = 10, \sigma = 10, \beta = 8/3$.

confidence is used. Even though a low dimensional causal inference problem, the Lorenz system is considered a challenging problem because of the nonlinearity of the relationships as well as the cancellation of causal effects when the system is considered on average. There exist other causal inference methods that are more suited to identify causal effects of a nonlinear system such as the Lorenz system. Specifically, nonlinear state-space methods such as convergent cross mapping [217]. However, these methods are less suited for time series of stochastic nature [9].

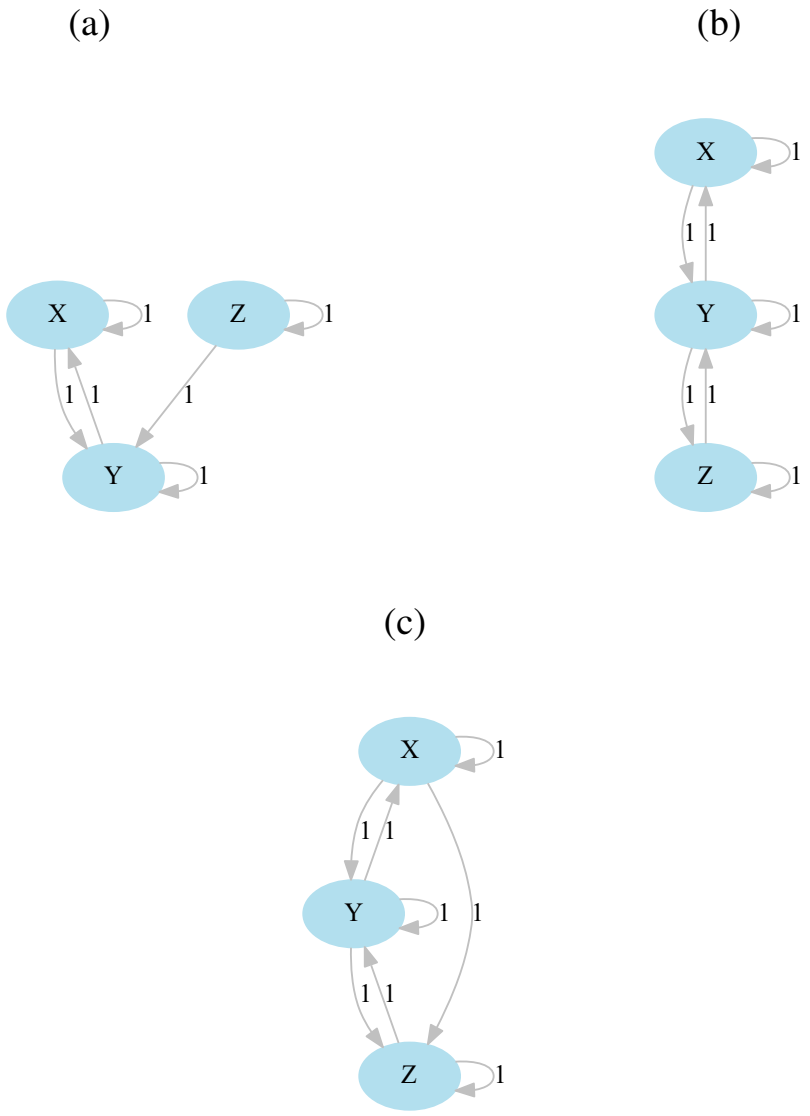


Figure C7: Summary of Interactions identified by PC stable. Here we present $1\delta t$ as 1. Parameters $\rho = 28, \sigma = 10, \beta = 8/3$. Results for a) $\alpha = 0.001$, b) $\alpha = 0.05$ and c) $\alpha = 0.2$

Even though causal network learning algorithms that use nonparametric tests to assess conditional independencies will allow to identify both linear and nonlinear causal interactions, the associated computational complexity make these methods unsuitable for a high dimensional problem as the one presented in Chapter 5. Therefore, we resort to this simpler approach as a starting point.