SPATIOTEMPORAL SCALES IN MODELING: IDENTIFYING TARGET SYSTEMS

by

Marina Baldissera Pacchetti

B.Sc. in Mathematics and Physics, University College London, 2009

M.Sc. in History and Philosophy of Science, London School of Economics and Political Science, 2010

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Marina Baldissera Pacchetti

It was defended on

June 22, 2018

and approved by

Robert Batterman, Professor and Chair, Department of Philosophy

Sandra D. Mitchell, Distinguished Professor, Department of History and Philosophy of Science

James F. Woodward, Distinguished Professor, Department of History and Philosophy of Science

Mark Wilson, Distinguished Professor, Department of Philosophy

Charlotte Werndl, Professor, Department of Philosophy (KGW), University of Salzburg

Dissertation Advisor: Robert Batterman, Professor and Chair, Department of Philosophy

SPATIOTEMPORAL SCALES IN MODELING: IDENTIFYING TARGET SYSTEMS

Marina Baldissera Pacchetti, PhD

University of Pittsburgh, 2018

My dissertation addresses neglected roles of idealization and abstraction in scientific modeling. Current debates about epistemic issues in modeling presuppose that a model in question uncontroversially represents a particular target system. A standard line of argument is that we can gain knowledge of a target system simply by noting what aspects of the target are veridically represented in the model. But this misses epistemically important aspects of modeling. I examine how scientists identify certain phenomena as target systems in their models. Building on the distinction between data and phenomena introduced by Bogen and Woodward, I analyze how scientists target systems from data and from basic theoretical principles. I show that there are two crucial empirical assumptions that are involved in identifying phenomena. These assumptions concern the conditions under which phenomena can be indexed to a particular length or time scale and the conditions under which one can treat phenomena occurring at different length or time scales as distinct. The role of these assumptions in modeling provides the basis for a new argument that shows how, in many cases, idealizations and abstractions in models are essential for providing knowledge about the world in so far as they isolate relevant components of a phenomenon from irrelevant ones. My analysis of the identification of phenomena also shows that structural

uncertainty arises in models when the scale of a phenomenon of interest is not properly identified. This clarification promises to improve the communication of the limitation of current climate models to policy makers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

My interest in idealizations in climate science was inspired by the work of Bob Batterman and Mark Wilson, to which I was first exposed during my time at Pitt. Their attention to the details of scientific practice in their discussion of epistemic issues in scientific modeling has inspired my research since I first encountered their work. There are several more people I am indebted to for helping me lay bricks that make up this dissertation.

First of all, I would like to thank my advisor Bob Batterman and committee members Sandy Mitchell, Charlotte Werndl, Mark Wilson and Jim Woodward. Their support and feedback have been stimulating and essential for the development of the main ideas presented in this dissertation.

There are no words to express my gratitude for Bob's guidance throughout the dissertation writing process. In addition to the many conversations we have had on the philosophy of science and the role of spatiotemporal scales in physics, his constant encouragement and endless patience in dealing with the many drafts he has read have been the main source of motivation throughout the years I have spent at Pittsburgh. I feel extremely lucky to have had him as my advisor and would not want to repeat this process with anyone else.

I am indebted to the graduate student community of HPS at the University of Pittsburgh. Their help has come especially in the form of feedback at the "Work in Progress" (aka WIPs) sessions at which I have presented chapters 2 and 4 of this dissertation. I would especially like to thank Haixin Dang, Josh Eisenthal and Gus Law, who have assisted me in keeping up with the writing and the challenges related to every stage of my graduate career at Pitt. Thank you BLED! Past HPS graduate students have also assisted me in various ways. Leif Hancox-Li and Catherine Stinson were especially helpful during the last year of graduate school.

I also thank Porter Williams and Michael Miller for their feedback on several drafts of papers that appear as chapters of this dissertation. Chapter 4, in particular, would not exist in the present form without Porter's patience and help. I also thank Erik Curiel for providing me with much needed "reality checks" and teaching me interesting physics and philosophy on a regular basis. Yoichi Ishida has been instrumental for the final editing of the dissertation. His attention to detail and form vastly exceeds mine.

I am also grateful for the support of my family and friends. In particular I would like to thank Michal De-Medonsa, Emanuela Baldissera Pacchetti and my parents for their unconditional support throughout the years.

This dissertation is dedicated to Erik Curiel: it's his fault I ended up writing a doctoral dissertation in the first place.

1.0 INTRODUCTION

Giere (1988), Cartwright (1999), Morrison and Morgan (1999), Wimsatt (2007), Pincock (2012), Weisberg (2013) have focused on the following question: Given that models are false of the world, how can they produce reliable knowledge? More specifically, how can models that necessarily idealize or abstract from their target systems provide reliable information about these systems? Attempts to answer this question usually focus on trying to establish a certain kind of relation between the model and the target system. A standard line of argument is that we can gain knowledge of a target system simply by noting what aspects of the target are veridically represented in the model. For example, early responses to the issue of idealization and abstraction relied on the notion of "Galilean Idealization" (McMullin 1985, Laymon 1985), according to which an idealization is initially introduced in the model but can be subsequently dispensed with, when more detailed models are constructed. [1] Some philosophers, especially Batterman (2009), have resisted this view of idealization and have argued that idealizations are in many cases indispensable for models to provide knowledge. Weisberg (2013) suggests that there are several non-competing ways in which idealizations are epistemically useful.

[1] McMullin's definition of idealization includes any simplification of something complicated (1985, 248), so it includes what has been more recently called idealization and abstraction (Jones 2005).

My dissertation is inspired, in part, by a lacuna in the literature on the role of idealization and abstraction in scientific modeling described above. Current debates about epistemic issues in modeling usually assumes that *what the model is a model of* is uncontroversial, or that it is beyond the scope of philosophical analysis. One consequence of the assumption that what a model is a model of is uncontroversial is that abstractions and idealizations are analyzed as constituent parts of the model that either are essential or subsequently need to be removed. My work takes a radically different approach, as the standard line misses important aspects of the epistemic role of modeling. I examine the processes by which phenomena are identified as targets.

Building on the distinction between data and phenomena introduced by Bogen and Woodward (1988), I discuss the details of how scientists identify relevant aspects of the world to be modeled both from data and from basic theoretical principles. My approach analyzes abstractions and idealizations as assumptions that are part of the process by which scientists obtain models of phenomena. Here, phenomena are stable, recurrent features of the world. By taking this different approach, I can explore the role of the assumptions that are crucial for identifying targets as phenomena and highlight the epistemic and ontological implications of these assumptions. I show that there are two crucial empirical assumptions that are involved in identifying phenomena. These assumptions concern the conditions under which these phenomena can be indexed to a particular length and time scale, and the conditions under which one can treat phenomena occurring at different length or time scales as distinct. The role of these assumptions in modeling provides the basis for a new argument that shows how, in many cases, idealizations and abstractions in models are essential for providing knowledge about the world in so far as they isolate relevant components of a phenomenon from irrelevant ones.

My analysis of the role of scale related assumptions in the context of identifying phenomena also allows me to clarify a concept that is part of the growing philosophical literature on epistemic issues in climate science. This is the concept of structural uncertainty. Typically, this is considered to be uncertainty about the structure of the equations that represent the climate.[2] Current philosophical analyses have identified an important problem, but they miss some important points. For example, they also do not consider the issue of identifying target systems, and as such their accounts cannot help in the mitigation of structural uncertainty; nor can they provide for clear demonstrations of the limitations of current climate models to policy makers. My analysis of the identification of phenomena removes some of this problematic vagueness. I show that one way in which structural uncertainty arises in models is when the scale of a phenomenon of interest is not properly identified. The scale related assumptions are in fact a tool for determining what components are relevant for adequately identifying and modeling phenomena of interest. Failure to properly use this tool introduces uncertainty in whether scientists are adequately identifying and modeling these phenomena. This characterization provides a way to individuate where and to what degree structural uncertainty is tied to misidentification of the scales at which phenomena are present. The analysis also provides insight into how to mitigate against structural uncertainty.

One of the most important aspects of target identification in science is the use of scales. Scales are understood as spatiotemporal scales, and these are fundamental for the definition of physical system. As Emanuel says, “A system must somehow be distinct in space and time and the transfer of physically relevant quantities across the boundaries of the system must be understood”

[2] See, for example, Parker (2006, 2010, 2011), Frigg et al. (2013, 2014), Stainforth, Allen, et al. (2007), Knutti (2008).

(Emanuel 1986, 6). However straightforward and obvious this statement might sound, it hides both ontological and epistemic subtleties in the atmospheric and climate sciences, and in science in general. Like climate scientists, population ecologists also recognize the difficulty of finding target systems, and the role that spatiotemporal scales play in identifying them. Finding the spatiotemporal scale at which recurring features can be observed is conducive to successful model building (whatever the purpose) (Levin 1992). Individuating the scales at which phenomena occur and modeling the dynamics of the phenomenon with the relevant variables and parameters is a central yet difficult enterprise in meteorology (Emanuel 1986).

Three different examples of the problem of scale for modeling phenomena illustrate the importance of this issue. These examples are derived from population biology (Levin 1992), oceanography (Stommel 1963), atmospheric science (Emanuel 1986), and show different but overlapping aspects of the problem of scale. The examples I provide raise important themes that will be discussed in this dissertation. These themes are the problem of identification of target systems, the assumptions used to identify the targets, their nature, and the role of spatiotemporal scales. In the rest of this chapter, I will briefly discuss the examples to show the centrality of the problem of scale.

1.1 ECOLOGY

Levin (1992) says:

> Our efforts to develop theories of the way ecosystems or communities are organized must revolve around attempts to discover patterns that can be quantified within systems and compared across systems. (Levin 1992, 1947)

Levin's point is that quantitatively precise descriptions of systems must capture the patterns within the systems of interest. Different systems will be described in terms of different patterns. The variables that represent physical parameters of interest quantify these patterns and define the target system. The physical parameters of interest are organized in terms of spatiotemporal scales. The difficulty is the quantification of the variability—the frequency and amplitude of a pattern in a data series—of a physical quantity. Quantifying the variability, in this case, involves identifying the aspects of the data series that are associated with the system. One important step in identifying these aspects is to find the relevant scale as a system, and to determine how the system interacts with other scales across its boundaries.

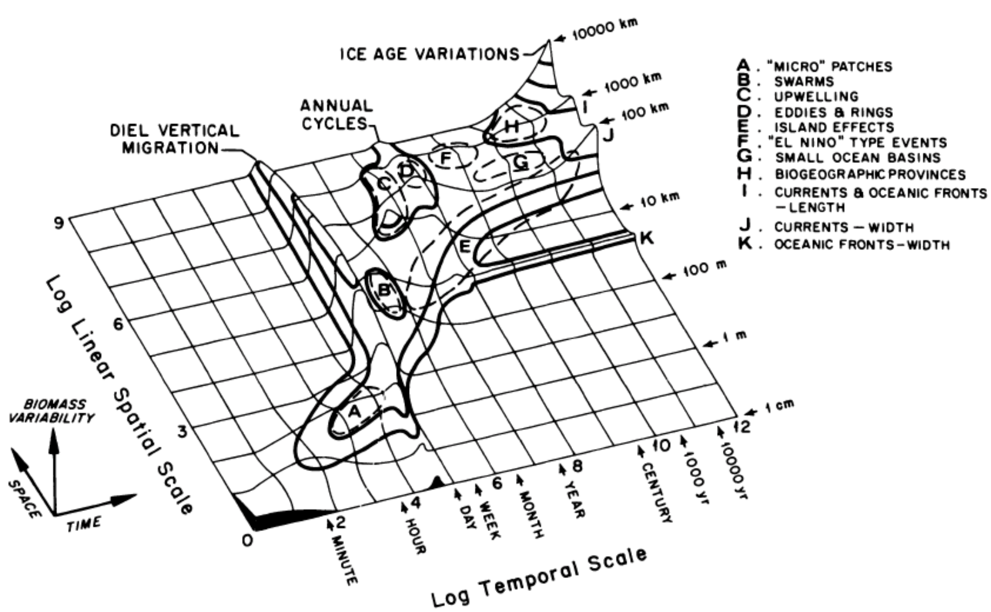

Figure 1.1. Spatial and temporal scales of zooplankton biomass variability. Reprinted with permission from Springer Customer Service Centre GmbH: Springer Nature, Patterns and Processes in the Time-Space Scales of Plankton Distributions, Haury et al., © (1978).

Figure 1.1 shows how a physical quantity of interest for population ecologists—biomass variability—organizes in space and time. Even if the biomass of plankton can have small variations at all scales, there are clear peaks in the graph at the characteristic spatiotemporal scale of processes of interest. The challenges in modeling these processes lie in having an observation network that accurately collects data series at regular spatial and temporal intervals and building a mathematical model that accurately represents the relevant parameters for the processes.

Levin recognizes that characterizations of phenomena are limited by our observational capabilities. He suggests that "our perception of events provides us with only a low-dimensional slice through a high-dimensional cake" (Levin 1992, 1945), where the "perception" is organized in space and time. Spatiotemporal scales are "imposed on us by our perceptual capabilities, or by technological or logistical constraints" (Levin 1992, 1945). Perceptual capabilities are aided by instruments, such as observational networks and the methods used to extract the relevant signal from the observations (such as Fourier analysis). These capabilities, however, can be compromised by lack of technology or lack of good strategies for setting up effective networks.

Levin also recognizes that the resolution of the variability of a system of interest depends on the scale at which the system is observed (Levin 1992, 1945). This dependency implies that a system and its dynamics can be described successfully if the right spatiotemporal scale at which the variability drives the system is recognized. The variability of the system can be thought in terms of its dynamics, i.e. those quantities that are part of the system that change in time. Which parameter is relevant (kinetic energy, biomass variability, etc.) depends in part on the modeling aims of the scientist.[3] Nevertheless, Levin claims that for any variability of interest, there is going to be a spatiotemporal scale at which variables and parameters can describe the relevant drivers of the variability:

---

[3] There are both empirical and normative considerations that are taken into account when identifying target systems. I discuss these issues in chapter 2.

> This is the principal technique of scientific inquiry: by changing the scale of description, we move from unpredictable, unrepeatable individual cases to collections of cases whose behavior is regular enough to allow generalizations to be made. (Levin 1992, 1947)

Levin sees a close connection between successful modeling and finding regularities at different scales such that they can be predicted and applied to similar instances. The technological and logistical constraints just mentioned limit the ability to recognize such scales.

Levin's view can be summarized as follows. There are recurrent patterns that can be recognized at different spatiotemporal scales. These are what scientists can identify as systems or phenomena. Modeling strategies must capture these patterns at the right scale, in order to quantify the variability (dynamics) so that predictions can be made and describe the dynamics so that the general physics (or biology) of it can be understood.

The ecologist Holling provides another example of the importance of scale for the definition of physical system:

> All terrestrial ecosystems are controlled and organized by a small set of key plant, animal, and abiotic processes. They form interacting clusters of relationships, each of which determines the temporal and spatial structure over a constrained range of scales. The overall extent of these influences covers at least centimeters to hundreds of kilometers in space and months to centuries in time. (Holling 1992, 451)

Holling claims that we observe patterns that arise in ecosystems. These patterns arise at characteristic spatiotemporal scales, at which we can identify key interacting relationships (in Holling's case control and organization of the components).

1.2 OCEANOGRAPHY

In a landmark paper published in 1963, the oceanographer Henry Stommel discusses the importance of an organized and systematic network of observations for studying ocean dynamics. He claims that different grid sizes of observation networks are needed to detect different phenomena. Stommel writes:

> [A] single net does not catch fish of all sizes; the existing net of tide-gauge stations does not suffice for a study of turbulence. It is necessary to decide which part of the spectrum of each variable one wants to measure. … Each plan must provide a definite significance level within a limited part of the spectrum *despite contamination from other parts of the spectrum*. (Stommel 1963, 573; my emphasis)

There are several important points. Grids of observational networks of different sizes are needed for observing different phenomena. Certain phenomena, such as turbulence, need a finer net, as it is a process that occurs on small to large spatiotemporal scales. The fact that processes occur at different spatiotemporal scales implies that different observational networks can and do detect stable recurring signals at many different scales, despite the fact that all scales are connected to each other.

The difficulty that both the observational scientist and the modeler need to take into account is the following. The network of observation stations needs to be on a grid size that allows for a spectrum that actually detects the signal of processes occurring at the scale of the grid size. Subsequently one needs to be able to detect the signal of the process in the time series of data collected by the net.

These points have both ontological and epistemic consequences. The first point is ontological: phenomena are characterized as existing at different spatiotemporal scales. The epistemic consequences are that observational networks need to have different grid sizes in order to be able to detect different phenomena, and that even at the relevant scale for a certain phenomenon, the signal from the quantity of interest that is measured needs to be distinguished from the noise coming from phenomena at other scales.

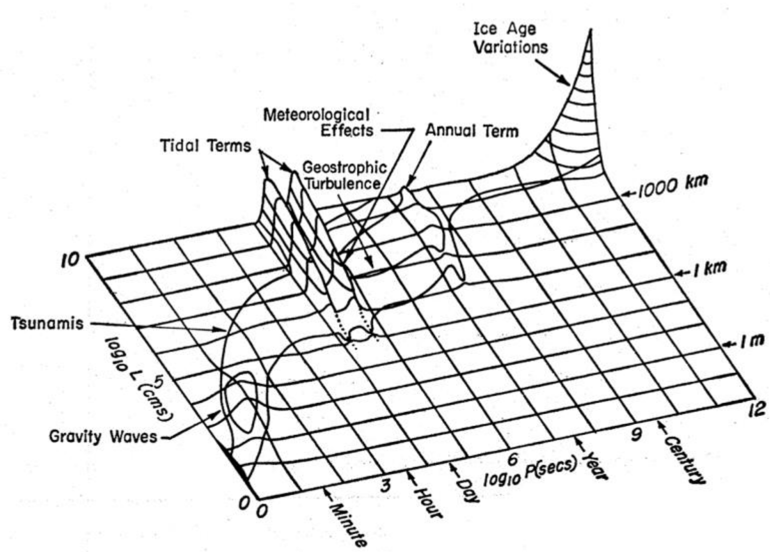

Figure 1.2. Spatiotemporal scales of processes occurring on the ocean surface. From Stommel (1963). Reprinted with permission from AAS.

Figure 1.2 provides an intuitive visualization of Stommel's points. In the graph, the two horizontal axes are space and time, while the vertical axis represents energy. The graph shows peaks of kinetic energy at different spatiotemporal scales, and these peaks represent different processes. Tides, for example, are processes that can be detected at temporal scales just above the hour and just below a day, and on spatial scales between one and one thousand kilometers. This means that an observational network that extends above these scales is not going to capture this phenomenon, and neither will an observational network that detects changes in movements of the ocean below one hour and on spatial scales below the kilometer.

Further, note that the graph has lines that connects the energy peaks, meaning that these peaks are interconnected. This implies that there is always going to be some interaction among the scales at which peaks are found, and the difficulty in building a successful model is two-fold. First, at the scale at which phenomena can be detected, one needs to distinguish between the variables that are relevant for the recurrence of the phenomenon of interest, and the noise from other scales needs to be left out. Second, the modeler also needs to be aware of interactions *across* scales, as these might sometimes be important for the dynamics of a model that intends to include different scales of motion.

1.3 ATMOSPHERIC DYNAMICS

While Stommel provides a primarily data-driven account of the importance of spatiotemporal scales in modeling, the atmospheric scientist Emanuel provides a general perspective on the problem of scales in meteorology.

As mentioned above, the location of a system in space and time, and the understanding of the variables that are relevant for the conservation within and the transfer across the system's boundaries is fundamental for the recognition and definition of a target system. The location of a target system is at the heart of the modeling practice, and it presents many challenges, some of which we have seen in the previous subsections. The atmosphere is a particularly interesting case study, as it is modeled as a fluid the energy spectrum of which is "smooth and continuous between the limits imposed by the mean free path of molecules on the short scale and the circumference of the earth on the large" (Emanuel 1986, 1). Nevertheless, phenomena such as cumulus clouds, hurricanes, weather fronts et cetera, are seen to occur and recur on many different but characteristic spatiotemporal scales—they can be seen as phenomena in their own respect that are relatively isolated from their environment. The challenge is to individuate systems at these characteristic scales (and the exceptions), construct models that capture the dynamics of the phenomena observed, and explore the way these phenomena transfer energy across their boundaries to other smaller or larger spatiotemporal scales.

The individuation of phenomena can be challenging. This challenge is due to a lack of consensus about the identification of processes and their characteristic scales, or about whether any such processes or spatiotemporal scales exist. According to Emanuel, this issue is particularly evident in the case of the so-called mesoscale[4]:

> A serious question confronting the mesoscale meteorologist is whether there are really inherent scales in the atmosphere that one might reasonably use to define the mesoscale. . . . Do there exist ordered processes in the atmosphere that generate kinetic energy on scales within Ligda's mesoscale domain (does a natural mesoscale exist), or does the "mesoscale" really consist only of a smooth, continuous, and uninteresting spectrum of disordered motions . . . ?" (Emanuel 1986, 5)

Ligda is a meteorologist who in the 1950s used the radar to investigate structures in the atmosphere and found stable structures recurring at scales smaller than the large scale[5] but larger than the small

[4] The mesoscale encompasses phenomena that are larger than a cumulus cloud but smaller than hurricanes.

[5] This is the scale where the dominant (fluid) motions are driven by the rotation of the earth and its circumference, as I will show below.

scale. He coined this range of scales the "mesoscale" and, since then, atmospheric scientists have attempted to find spatiotemporal scales at which kinetic energy is generated in this range and the phenomena associated therewith.[6]

The scientific issues tied to identifying target systems mentioned in the cases described above will drive the philosophical arguments in this dissertation. Assumptions about the spatiotemporal scales at which phenomena occur play an important role in identifying target systems, and as such these assumptions can provide insight into the debate on the role of idealizations and abstractions in modeling. The use of these assumptions is important both in the process of collecting and analyzing data, and in the process of deriving appropriate mathematical models. The philosophical analysis of classical and novel philosophical problems such as idealization and structural uncertainty I provide shows that the approach taken in this dissertation is an approach that can provide a new perspective on these debates. While it is true that the identification of targets is part of the creative process of science, philosophers can identify those rational constraints that make the scientific process a reliable source of knowledge. Examples of the success of such work can be found in Bogen and Woodward (1988, 1992, 2005), Woodward (1989, 2000, 2011), and Peschard (2012a, 2012b).

The rest of the dissertation is organized as follows. Chapter 2 takes up Bogen and Woodward's distinction between data and phenomena. This distinction raises the need to understand the structure of the data-to-phenomena and theory-to-phenomena inferences. I suggest that one way to study the structure of these inferences is to analyze the role of the assumptions involved in the inferences: What kind of assumptions are they? How do these assumptions

---

[6] The generation of kinetic energy is associated with a dynamic process that can be identified at that scale that generates this energy.

contribute to the practice of identifying phenomena? In this chapter, using examples from atmospheric dynamics, I develop an account of the practice of identifying the target in the data-to-phenomena and theory-to-phenomena inferences in which assumptions about spatiotemporal scales play a central role in the identification of parameters that describe the target system. I also argue that these assumptions are not only empirical but they are also idealizing and abstracting. I conclude the chapter with a reflection on the role of idealizations in modeling.

Chapter 3 analyzes in more depth the way that these scale related assumptions are employed in the data-to-phenomena and theory-to-phenomena inferences. The identification of phenomena for the case of data-to-phenomena inference is studied for the case of El Niño. While the details of the physics responsible for this phenomenon are still debated today, a major step in the scientific understanding of the phenomenon occurred in the nineties, when a new observation network was developed in the Pacific basin. The considerations that guided the development of this observation network were similar to the ones mentioned by Stommel described above: identifying physically meaningful variabilities in the parameters of interest involves making substantial assumptions about the spatiotemporal scales at which these variabilities are thought to occur.

The identification of phenomena for the case of the theory-to-phenomena inference is studied by analyzing an important parametrization scheme in climate models. This parametrization scheme is an important tool for suppressing irrelevant details and making the models more computationally manageable, and, more importantly, helping scientists understand the behavior of target systems. Here, again, scale related assumptions play an important part in determining when this kind of parametrizations are justified and can indeed provide insight into the physical processes responsible for the phenomenon. The analysis of this parametrization scheme is a further exemplification of the argument made in chapter 2: the scale related assumption employed in the justification of this parametrization is at the basis of the selection of parameters used to describe the phenomenon and is also an idealizing and abstracting assumption.

Chapter 4 analyses the concept of “structural uncertainty” from the perspective of target system identification. Structural uncertainty is an epistemological problem climate scientists and philosophers have been concerned with especially in the context of climate science. This type of uncertainty is defined as uncertainty about whether a mathematical model accurately represents its intended target. A clear account of structural uncertainty is essential for the effective communication of the science to policy makers. I identify three desiderata that an account of structural uncertainty should meet in order to be useful to philosophers, scientists and policy makers. After reviewing some of the most influential accounts of structural uncertainty and addressing their shortcomings, I argue that a useful account of uncertainty in climate science can be given by looking at how target systems are identified in the context of climate modeling. In particular, I show how the use of assumptions about spatial and temporal scales that are involved in identifying and modeling target phenomena can lead to structural uncertainty.

Chapter 5 analyses the ontological implications of the use of scale related assumptions. I suggest that the use of these assumptions provides a different perspective on the current literature on “levels.” I describe three of the prominent accounts of levels in the philosophical literature. These appear in the works of Levins, Wimsatt, and Mitchell. These accounts offer different philosophical perspectives on levels: Levins provides a descriptive epistemological approach, Wimsatt offers an ontological account of levels, and Mitchell uses the concept of level as a tool in her philosophical analysis of epistemic issues and metaphysical considerations of modeling a complex world. Their accounts overlap in many respects. In light of the discussion of identifying target systems in various branches of science given in the previous chapters, I suggest that the concept of level is best understood in terms of relevant variables and parameters that capture recurrent patterns occurring at particular spatiotemporal scales.

2.0 A ROLE FOR SPATIOTEMPORAL SCALES IN MODELING

2.1 INTRODUCTION

Among the many philosophically interesting topics related to modeling, one topic that has not received much attention is how scientists determine their target systems in the world.[7] As mentioned in chapter 1, two issues in the modeling literature bear on this topic. One concerns the role of idealization and abstraction in modeling, and the other concerns the relationship between theory, data and phenomena. This chapter shows how focusing on the role of scale related assumptions for the latter of these issues can inform the former.

In a series of papers, Bogen and Woodward have developed an influential account addressing the issue of the relationship between theory, data and phenomena (Bogen and Woodward 1988, 1992, 2005; Woodward 1989, 2000, 2011). Their account involves three important points. First, theory does not explain data (Bogen and Woodward 1988, 305) and there are two conceptually distinct inferential processes in scientific practice, both aimed at characterizing features of the world of scientific interest: one in which explanations of phenomena are derived from theoretical principles, and one in which phenomena are inferred from data (Woodward 2011, 168). Second, inferences from data to phenomena are ampliative and usually

[7] Two exceptions are Elliot-Graves (2014), which investigates the ontological status of target systems, and Peschard's (2012a, 2012b) work discussed below.

involve empirical assumptions. The inferences are ampliative insofar as they "go beyond the data" (Woodward 2011, 173), and the assumptions are empirical in so far as they can be either true or false (Woodward 2011, 173).[8] Third, pragmatic considerations of the scientist, such as her research interests and resources, can also play a role in these inferences (Woodward 2011, 174).

Bogen and Woodward's account thus raises the need to understand the structure of the data-to-phenomena and theory-to-phenomena inferences (see also Woodward 2011, 170). One way to study the structure of these inferences is to analyze the role of the assumptions involved in the inferences: What kind of assumptions are they? How do these assumptions contribute to the practice of identifying phenomena? In this chapter I develop an account of the practice of identifying the target in the data-to-phenomena and theory-to-phenomena inferences. This account relies on recognizing the important role of assumptions about spatiotemporal scales in the identification of parameters that describe the target system. I also argue that these scale related assumptions are not only empirical, but they are also idealizing and abstracting. In this chapter I will supplement my argument with brief examples from atmospheric science. In the next chapter, I will provide more detailed examples from both climate science and oceanography.

The atmosphere is a particularly useful case study for illustrating the role of scale related assumptions in the process of identifying target systems. As mentioned in chapter 1, one of the main problems in meteorology is to discretize the continuous energy spectrum by individuating the phenomena that are associated with characteristic peaks in this spectrum. In order to model the dynamics of the phenomena observed, and explore the way these phenomena transfer energy across their boundaries to other smaller or larger spatiotemporal scales, scientists need to make

[8] Woodward also suggests that these assumptions can be theoretical when they are about factors that cannot be observed (Woodward 2011, 173).

assumptions about the extent to which these phenomena can be separated from phenomena at other scales.

To delve deeper into the challenges involved with identifying phenomena, it is worth revisiting the important passages of Emanuel's discussion quoted in chapter 1, repeated here for easy reference:

> A serious question . . . is whether there are really inherent scales in the atmosphere that one might reasonably use to define the mesoscale . . . Do there exist ordered processes in the atmosphere that generate kinetic energy on scales within Ligda's mesoscale domain (does a natural mesoscale exist), or does the "mesoscale" really consist only of a smooth, continuous, and uninteresting spectrum of disordered motions . . . ? (Emanuel 1986, 5)

The problem that the meteorologist has to face is whether one can resolve phenomena within the range of the mesoscale, i.e., whether there are stable patterns that can be characterized as phenomena. Emanuel's quote illustrates that identifying phenomena is not a trivial endeavor, since scientists need to be able to distinguish what, if at all, is a genuine signal of an ordered motion from the noise of disordered motions. There are several challenges associated with identifying a genuine signal: first of all, the observations need to occur at a particular scale. Second, the data needs to be interpreted, for example by choosing the correct tools to eliminate noise. Third, the interpretation needs to be justified—the lack of a theoretical justification can lead scientists and philosophers to discard an interpretation as purely conventional or pragmatic. Major assumptions in these processes are scale related: phenomena are assumed to occur at particular spatiotemporal scales, and the scale separation assumption states that if the scales at which variables recur are sufficiently different from one another, then they describe different phenomena. For example, the scale at which cumulus clouds occur is much smaller than the scales at which large scale phenomena, such as cyclones, occur. It is largely in virtue of this scale separation that they are considered different phenomena. The problem of identifying a target system is therefore tied to finding regularities occurring at various spatiotemporal scales and identifying what components describe the system's dynamics.

I proceed as follows. In section 2.2 I use Bogen and Woodward's distinction between data, phenomena and theory to describe the process of identifying target systems and I show that target system identification involves two important scale-related assumptions: the scale existence and the scale separation assumptions. I support my account with examples from the atmospheric sciences. I also explain how these assumptions are both idealizing and abstracting. Section 2.3 further clarifies my account by addressing three objections to my account. These are McAllister's (1997) objection concerning the arbitrariness of certain assumptions in the data-to-phenomena inference, a possible circularity in the data-to-phenomena and theory-to-phenomena inferences, and the importance of what Peschard (2012a, 2012b) calls non-empirical "relevance judgments" (Peschard 2012b, 749) in the identification of phenomena. Section 2.4 draws some conclusions about the nature of idealization in the context of the current philosophical debate given the discussion in sections 2.2 and 2.3.

2.2 IDEALIZATIONS AND INFERENCES TO PHENOMENA

In Bogen and Woodward's framework, the main difference between data and phenomena is the following. Phenomena repeat themselves under many different conditions: they have "stable, repeatable characteristics" (Bogen and Woodward 1988, 317), and can still be recognized as such despite these different conditions. Data, on the other hand, are tainted by the idiosyncrasies of the way the data are collected (Bogen and Woodward 1988, 317). In other words, data contain both the signal of the phenomenon of interest and the noise coming from possibly irrelevant factors of the world, while phenomena themselves abstract away from these irrelevant factors. Despite the fact that phenomena are abstracted from the irrelevant details generated by the measurement process, Bogen and Woodward claim that phenomena belong "to the natural order itself and not just to the way we talk about or conceptualize that world" (Bogen and Woodward 1988, 321).[9]

This characterization of data and phenomena lies at the core of Woodward's claim that there are two different inferences to phenomena in scientific practice. First, theory explains phenomena and not individual pieces of data because the data contains both the signal of the phenomenon and various idiosyncrasies from measurement (Woodward 2011, 166). For example, scientists can explain the quantitative value of the melting point of lead by invoking characteristics of electron bonds and the presence of delocalized electrons present in lead. This is what Woodward calls a systematic explanation of a phenomenon from theory (Woodward 2011, 166). Second, a collection of data that measures the melting point of lead will be of statistical nature, the measurement being influenced by other factors the details of which are not typically known by the scientist (Woodward 2011, 167). The theory-to-phenomena and data-to-phenomena inferences can

[9] A possible worry is that what counts as phenomena may be interest-relative. This would imply that there is no principled distinction between phenomena and data as Bogen and Woodward envisage. To address this worry, it may be useful to distinguish between what counts as phenomena and what counts as explananda. According to Bogen and Woodward's view, phenomena are defined as regularities in the world. Of course, scientists may be interested in explaining the absence of regularities, such as the absence of a tide in a particular location. Thus, what counts as explananda may be interest relative, and not all explananda have to be phenomena in Bogen and Woodward's sense.

be more or less independent of each other, especially when the theoretical apparatus is not well developed (Woodward 2011, 170).

Insofar as theory-to-phenomena and data-to-phenomena inferences are ampliative, they involve certain assumptions in addition to theoretical principles and data for reaching a conclusion. Thus, these assumptions are crucial for correctly identifying phenomena, and their role should be investigated further. Woodward suggests that these assumptions may vary greatly from case to case, but there are also generic assumptions, such as the assumption about the probability distribution governing the error in a data series (Woodward 2011, 172–173). In the next subsections, I show how two such general assumptions are scale related and play a central role in the process of identifying dynamic phenomena.

2.2.1 Assumptions About Scales

Two crucial assumptions involved in the identification of target systems are about the spatiotemporal scale at which the phenomenon occurs. These assumptions are:

> 1. *The scale existence assumption*, which says that phenomena exist at characteristic spatiotemporal scales; and
>
> 2. *The scale separation assumption*, which says that if two patterns are sufficiently separated in space and time, then they characterize two separable phenomena.

These assumptions isolate the target system from its environment, and they are idealizing and abstracting in the most commonly accepted senses of the terms, such as those given by Jones (2005): an idealization is a distortion of the system, a useful falsehood about the system. An abstraction leaves out features of the system (or the world) that are irrelevant for a useful or satisfactory description of the system (Jones 2005, 175). The scale existence and scale separation assumptions are idealizing in so far as they effectively remove variations in the data or terms in the equations that represent actual variabilities in the world. Smoothing of temperature fields (a consequence of the scale separation assumption), for example, introduces a distortion in the representation of the actual temperature fields found in the atmosphere (Emanuel 1986, 7). The same assumptions are also abstracting in so far as certain variations in the data are irrelevant for describing the target system at hand, and therefore can be removed. Smoothing the temperature fields at a particular scale allows for the isolation of variabilities in the fields at other, related, spatiotemporal scales.

In meteorology, identifying phenomena at their characteristic scales and modeling their dynamics using appropriate variables and parameters are central yet difficult enterprises (Emanuel 1986). These tasks involve locating a system in space and time, as well as identifying the variables relevant for describing the conservation of energy within, and the energy transfer across, the system's boundaries (Emanuel 1986, 6). This means that in order to characterize a target system, the scientist must both determine what parameters describe the system, and identify the scales at which these parameters recur. The correct choice of these parameters and their scales relies on the right application of the above assumptions.

Considerable challenges in modeling atmospheric phenomena are related to identifying the phenomenon and its components at their characteristic scales. These challenges are also related to the application of the above assumptions. Consider the case of cyclones: a cyclone can usually be distinguished as a phenomenon, yet it is not completely isolated from its environment (Emanuel 1986, 6). Further, despite the fact that some characteristics of cyclones can be clearly observed and identified at their characteristic scales, there are no models that can describe their dynamics accurately enough to make predictions (Montgomery and Smith 2014). The challenges for cyclone modeling derive from the multiscale nature of the phenomenon. The appropriate parameters to model the target system at the spatiotemporal scales at which it occurs ("the underlying physics") have not been isolated from the "unsystematic details of the individual cases" (Davis and Emanuel 1991, 1929).[10] In other words, neither theoretical principles nor data analyses converge to a model that represents the most relevant variables that describe the cyclone's dynamics.

The meteorologist Kerry Emanuel describes three strategies scientists have adopted to identify and describe target systems based on their characteristic scales: one empirical, one based on utility, and one theoretically motivated. The first and last of these strategies can inform the structure of the data-to-phenomena and theory-to-phenomena inferences. As I show in the next two subsections, these inferences importantly rely on the scale existence and scale separation assumptions.

2.2.2 Idealization and Data to Phenomena Inferences

In the data-to-phenomena inference, phenomena are identified from a collection of data. This inference involves distinguishing a signal, that is characteristic of the phenomenon, from noise, which may come from errors in measurement or other variations in the same parameter that are not characteristic of the phenomenon. The scale existence and separation assumptions enter this inferential process in the following way: the scale existence assumption is needed to focus on one recurrent pattern in the world. The scale separation assumption allows the scientist to isolate one

[10] This point is made also in Bogen and Woodward's discussion of "systematic explanations" (1988, 323). What Davis and Emanuel calls the "unsystematic details" are the data of quantities of interest of individual hurricanes the analysis of which does not lend itself to a unified interpretation.

pattern from another in order to differentiate between phenomena. By focusing on one pattern and distinguishing it from other patterns and other scales, modelers suppress irrelevant variabilities in a data series, abstracting these variabilities away. Further, these assumptions are invoked in data-driven strategies for identifying phenomena in terms of their characteristic scales of motion[11] both in the collection and interpretation of observations of physical quantities.

Phenomena can be identified through visual observations, weather maps, or spectral analysis. The main challenge for this identification of phenomena is to detect a pattern in the data at its characteristic scale despite noise from other scales. Weather maps, such as isobaric maps,[12] are one of the primary tools that exemplifies target system identification at scales that are not observable to the naked eye.[13] These maps heavily rely on the scale existence and separation assumptions both when data is collected and when it is analyzed for creating the map.

For example, when creating a map, scientists assume that a pattern can be detected at a particular scale (be this for experimental or theoretical purposes) in order to set up a data collection network, and subsequently make use of statistical analysis in order to decompose the data series into its various components.[14] If the components are separated enough in spatiotemporal scale, then one scale is—at least in the case of the instantaneous map—neglected in order to identify the target system of interest, introducing the idealization in the target system identification process. Spectral analysis of time series data relies similarly on these assumptions, and is aimed at providing

---

[11] Emanuel (1988, 1) uses the term "definition of scale", which should be understood as the identification of processes observed in the atmosphere in terms of their characteristic scales of motion.
[12] An isobar is a line of constant pressure on a map.
[13] Naked-eye observations are rather unproblematic. However, they only have a limited scope, as they are constrained by the visual field of an observer. For a discussion of the limitations of naked eye observations, see Houze (1993, Ch.1).
[14] The components do not need to be strictly periodic.

a rigorous quantification of the scales of phenomena identified by visual observations and/or maps. Spectral analysis usually aims at detecting the signal of the phenomena at their characteristic scale from the time series of observations. In this case, clearly separated peaks in the kinetic energy spectrum of a time series of data collected from the atmosphere or ocean surface indicate that there are stable recurrences of those energy values at a particular scale. These recurrences, in turn, correspond to a signal from phenomena recurring at characteristic scales.

Figure 2.1, reproduced below from chapter 1 for easy reference, shows some characteristic scales of motion associated with ocean surface processes, where the third, vertical axis is kinetic energy. In this case, kinetic energy is the parameter of interest. While the continuity of the spectrum indicates that there still is interaction across scales, the separated peaks in kinetic energy are associated with clearly distinguishable phenomena, like the tidal motion of the surface of the sea.

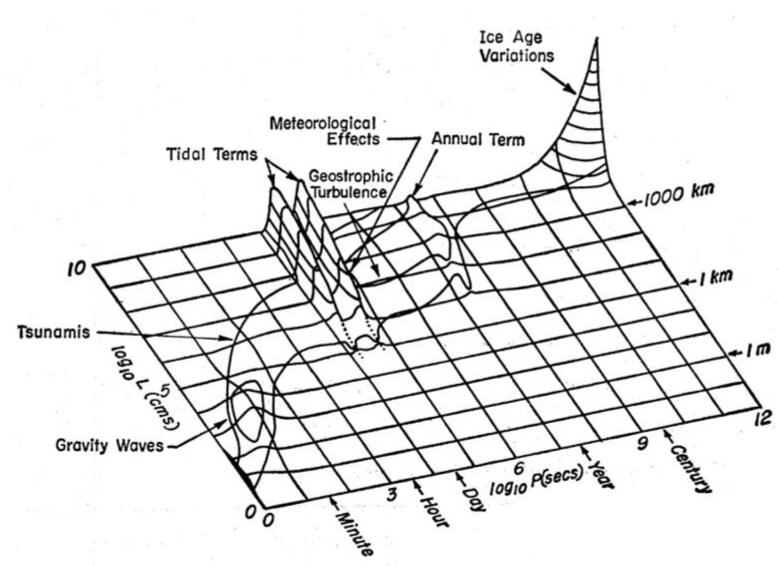

Figure 2.1. Spatiotemporal scales of processes occurring on the ocean surface. From Stommel (1963). Reprinted with permission from AAS.

The identification of phenomena and the successful application of the scale existence and separation assumptions depend on two main factors: the scale of the grid of observation stations and the data analysis. As a consequence, empirically driven identification of phenomena, and the correct application of the scale related assumptions, can fail in various ways. For example, a lack of rigorous systematization of the observation networks to study the dynamic processes in the ocean at different spatiotemporal scales may complicate statistical analysis. Further, a grid that resolves only one spatiotemporal scale is not going to capture all phenomena occurring in the ocean or the atmosphere (be this extremely coarse or fine-grained): different phenomena are observed using observational networks that resolve different spatiotemporal scales. For example, observing sea surface elevation only once a day (i.e., a temporal scale that captures variations from one day to another) does not capture the variation in sea surface elevation due to tidal motion. This is the case because peaks in kinetic energy due to tidal motions occur at and below the temporal scale of a day, as can be seen from figure 2.1. Such a time series, on the other hand, will capture variations due to geostrophic turbulence and meteorological effects, as these occur on scales slightly longer than the daily temporal scale. Note also that this time series will not be able to provide a clear characterization of larger scale variability, such as variations in sea surface elevation due to the melting and freezing of polar ice ("Ice Age Variations" in figure 2.1). Obtaining a trend for such a large-scale variability would involve averaging the data in a way that eliminates smaller scale variations, which is highly non-trivial.[15] These examples illustrate that the relation between phenomena and the spatiotemporal scales at which they occur is a central topic

[15] For example, scientists must make explicit assumptions about the distribution of smaller scale variability, and whether such small-scale variability is correlated with the scale of interest. This implies that knowledge of particular phenomena at one scale can inform how to extrapolate phenomena at other scales: the variability of some parameters at one scale can guide assumptions about how to average data at some contiguous scale.

for dynamic modeling of phenomena in the ocean and in the atmosphere.[16]

The scale related assumptions invoked in this case further illustrate the claim I made earlier that these assumptions are both abstracting and idealizing. As can be seen from figure 2.1, the parameter that is used to identify oceanic phenomena varies continuously from very small to very large spatiotemporal scales, and peaks occur at various, separated grip points. This feature has various implications. On the one hand, the scale existence and separation assumptions allow one to abstract the essential characteristics at which the phenomenon manifests itself—the scale at which the phenomenon recurs—from other, not directly relevant features: variations at other scales. The scale existence and separation assumption are also idealizing: they introduce a distortion in so far as they discretize a continuous spectrum.

2.2.3 Idealization and Theory to Phenomena Inferences

Like in the data-to-phenomena inference, the theory-to-phenomena inference also relies on the scale separation and scale existence assumptions. These assumptions are at the basis of scale analysis, an important component of the target system identification process. Scale analysis is further driven and substantiated by the interpretation of the physical processes that govern the phenomenon.[17] Hence, the existence of phenomena is inferred from dynamical and physical arguments and is based on the assumption that characteristic scales of motion exist, and that they can be separated from phenomena at other scales of motion. This inference proceeds in two main

[16] The atmosphere and the ocean are considered analogous, and they are studied with the similar observation techniques and the same equations, up to the values of constant parameters such as viscosity.

[17] This is where the line between empirical assumptions and normative relevance judgments gets blurrier. Peschard (2012a, 56) also implicitly acknowledges that there might be some ambiguity. I will discuss the relation of Peschard's work to mine in the next section.

stages.

First, possible scale related constraints on the dynamics of phenomena have to be identified. These are fairly independent of the details of the theory. For the case of atmospheric phenomena, these constraints are the sphericity and rotation of the earth, the depth of the atmosphere and its density. These are geometric boundaries that provide a first constraint on the possible scales of phenomena, and, as a consequence, on the application of the scale existence and separation assumptions. Second, physical arguments are used to justify the relevance of particular scales of motion for phenomena identified with these assumptions. In atmospheric dynamics, physical arguments are drawn from fluid dynamics: characteristic scales are described in terms of normal modes of oscillation and characteristic instabilities of the fluid.[18]

An example of a constraint on the target system identification in the first stage of the inference is the rotation of the earth, the most important geometrical constraint that qualitatively determines large scale phenomena in geophysical fluid dynamics. The so-called Rossby Number (RN) is a dimensionless parameter that allows one to quantitatively determine what terms in the equations of motion can be eliminated in order to describe the dynamics of large scale phenomena by means of dimensional analysis.[19] The RN, as other dimensionless parameters, is defined independently of the equations of motion. If L is the parameter for the characteristic scale of motion

[18] The following analysis is taken from Pedlosky's (1987) standard text.
[19] Dimensionless parameters are parameters that do not have a physical dimension (such as length, or velocity) attached to it. Dimensional analysis compares the relative size of dimensionless parameters when certain operations are performed, such as choosing a particular ordering relation. An ordering relation is a relation between dimensionless parameters. It allows one to study the behavior of one parameter when another parameter varies (this usually involves a limit), and to apply the scale separation assumption. I will provide an example of ordering relation later in the section.

of a phenomenon (the horizontal spatial variation of the dynamical fields of a phenomenon),[20] and U is the characteristic velocity scale (the velocity of the phenomenon as a whole),[21] then the RN is given by

$$\varepsilon = \frac{U}{2\Omega L}$$

where Ω is the period of rotation of the Earth. For rotation to play a role in the motion and structure of the system, ε must be of the order of one or less. The period of time it takes a fluid element to cross the whole system (the quantity L/U) has to be much smaller than the period of rotation of the earth for the rotation of the earth to have an effect on it. This consideration imposes some constraints on the possible spatiotemporal scales of large scale phenomena: values of ε and Ω, the geometrical boundary, restrict possible values of L/U. The RN therefore quantifies the geometrical constraint that the rotation of the earth imposes on the dynamics of phenomena. Since the RN is a function of the period of rotation of the earth, the stable patterns that arise at this scale are described in terms of a stability of the fluid that involves the Coriolis force, one of the inertial forces that drive the dynamics of motions that are affected by the earth's rotation. Since stabilities are described in terms of a balance of forces, the Coriolis force is going to play a dominant role in the description of large scale phenomena that arise from these balances - and, as a consequence, is going to be irrelevant for the dynamics of phenomena the RN of which is greater than one. Further, when the large-scale balance is in place, small deviations from this balance are irrelevant. This provides further physical justifications for the elimination of irrelevant terms in the equations.

In the second stage, the scale-related constraint comes from physical arguments. For large-

[20] Systems in the atmosphere are usually characterized by waves in a field of varying pressure, density, temperature, etc.
[21] The values of L and U are usually established observationally.

scale horizontal motions, the balance is described in terms of the geostrophic approximation.[22] The role of the geostrophic approximation is to eliminate the deviations from geostrophic motion, a motion in which the Coriolis force and the pressure gradient force are in balance. This balance is invoked to justify the elimination of the variations in the flow occurring at other scales, which are considered irrelevant for large scale phenomena. The equations of motion that are based on this balance yield solutions that describe the main motions of the large-scale atmosphere, such as Rossby waves.[23]

The above considerations are based on the scale existence assumption, as dimensional analysis is based on the assumption that "a single, well-defined scale for the velocity and its derivatives exists . . . [and] the magnitude of the terms in the equations of motion can be estimated in terms of these scales" (Pedlosky 1987, 340). This means that in order to find the simplified equations of motion that describe the behavior of a phenomenon at a particular scale, we need to start from the assumption that such a scale exists. From this assumption and the description of the relevant balances, scientists choose the ordering relation for the dimensionless parameters that are used in the equations of motion. The ordering relation allow them to determine that some terms in the equations are of a small enough order to be neglected, granting the application of the scale separation assumption.

The assumptions just described are abstracting and idealizing. Applying the assumption that phenomena occur at characteristic spatiotemporal scale abstracts away from the small variations in scale that can occur due to negligible processes at other scales. Further, the physical interpretation of the balances that characterize the scales of cyclonic and anticyclonic systems also

[22] Geostrophic balance is a balance in atmospheric wind flow between the pressure gradient force and the Coriolis force.
[23] For a full explanation and derivation of Rossby waves, see Pedlosky (1987, 374–385).

assumes that the vertical shear and stratification of the flow is constant. This introduces some idealizations in the description of the oscillations and instabilities that arise in the actual atmosphere (Emanuel 1986, 7). However, such simplifications are essential for the description of phenomena as these simplifications isolate them from scales of motion that are not directly relevant to the phenomenon.

2.2.4 Assumptions About Scales: Reprise

The above analysis highlights that the scale existence and separation assumptions play a crucial role in the observation, recognition and dynamic modeling of systems in the atmosphere. However, the application of the scale existence and scale separation assumption are only particularly effective when a scale can be clearly identified, either theoretically with the RN or empirically with data analysis. Consequently, one obstacle to identifying the scale of a phenomenon is the lack of clearly identifiable peaks in the parameters of interest.[24] In this case, the assumption about the existence of a characteristic scale can be unreliable.

To overcome this obstacle, scientists take advantage of the fact that theory-to-phenomena and data-to-phenomena inferences can be mutually reinforcing. The theoretical derivations can in fact provide a justification for the numerical relations between parameters obtained by observations, and vice versa. Woodward claims that "the data-based reasoning to P and the fact that T predicts P mutually reinforce our confidence that P is real—the two sets of consideration are in a positive feedback relation with each other" (Woodward 2011, 178). In sections 3.2 and 3.3, I show how the structure of the theory-to-phenomena and data-to-phenomena inferences

[24] This obstacle may be due either to the fact that a phenomenon occurs on many scales, and therefore a peak is unobservable, or to the fact that indeed there are no phenomena—no recurring stabilities—of interest at that scale.

substantiates Woodward's claim.

2.3 OBJECTIONS

So far, I have argued that the data-to-phenomena and theory-to-phenomena inferences involve empirical assumptions about the scale at which the phenomena occur and that these assumptions are a constraint imposed by the world on the identification of phenomena. Here I address three criticisms relevant to my account. One is McAllister's worry about the arbitrariness of certain assumptions involved in the data-to-phenomena inferences. The second is the potential circularity of the data-to-phenomena and the theory-to-phenomena inferences. The third is Peschard's idea that non-empirical assumptions rather than empirical ones play an important role in the identification of phenomena. I show that the analysis of the role of scale existence and scale separation assumptions in the process of identifying target systems can resolve these criticisms.

2.3.1 Arbitrariness

McAllister notes that there are infinitely many ways in which scientists can identify patterns in a data set (McAllister 1997, 220). However, only some of these patterns should count as phenomena in Bogen and Woodward's sense. Thus, McAllister asks how Bogen and Woodward might be able to answer the question of which phenomena are part of the natural order of the world and which patterns are noise (McAllister 1997, 222). He suggests that for Bogen and Woodward to answer this question, they would have to acknowledge that some patterns get counted as phenomena by stipulation (McAllister 1997, 221–222). In other words, the question is not settled entirely by empirical considerations.

McAllister is right in suggesting that there is some relativity to the interest of the scientist. For example, referring to figure 2.1, if a scientist is interested in “Geostrophic Turbulence,” then “Tidal Terms” will be noise in the dataset, and vice versa. Woodward replies to McAllister’s criticism by suggesting that assumptions involved in the scientist’s inferences are “empirical claims, that given a particular context, are either true or false” (Woodward 2011, 179). How the context of research is set up depends on the interest of the scientist, but in a given context, the assumptions used in the inferences to phenomena are subject to specific empirical tests. The account of the data-to-phenomena inferences developed in section 2.2 reinforces this response. Once the choice has been made about the scale at which data has to be collected, the application of the scale existence and scale separation assumption can either succeed or fail, depending on whether the scientists is collecting and analyzing the data at scales at which recurrent patterns occur in the world. As can be seen in figure 2.1, peaks in the energy spectrum correspond to recurrences of phenomena at particular spatiotemporal scales. This means that when data is collected and interpreted, there are scales at which phenomena in Bogen and Woodward’s sense can be identified as well as scales at which there are no phenomena to be resolved.[25] In other words, if scientists choose the ‘wrong’ scale at which to collect and analyze data, they are not going to be able to resolve the peaks in the spectrum that are associated with phenomena.[26]

It is true that when there are no clear peaks in the spectrum, as is the case with the mesoscale, scientists do face difficulties in identifying relevant scales of motion. Nevertheless, the fact that

[25] As noted in footnote 9, this of course does not mean that there will be nothing of explanatory interest.
[26] Levin also makes this point: “This is the principal technique of scientific inquiry: by changing the scale of description, we move from unpredictable, unrepeatable individual cases to collections of cases whose behavior is regular enough to allow generalizations to be made” (Levin 1992, 1947).

pragmatic choices do sometimes feature in data analysis does not imply that all phenomena are individuated exclusively by pragmatic choices. One seemingly pragmatic aspect of empirical identification of patterns in the data is the method of collection and interpretation of data. New observation techniques and tools for statistical analysis allow scientists to detect patterns at scales that were previously unobserved, as was the case for mesoscale phenomena: while before the 1950s atmospheric scientists were focusing only on the large scale, the introduction of radar as a tool for observation allowed scientists to collect new kinds of data that allowed for the resolution of patterns at scales that were previously unobservable. Similarly, the use of more sophisticated statistical tools allowed the meteorologist Sir Gilbert Walker to identify the "Southern Oscillation," a large scale climatic phenomenon (see Katz 2002, Pincock 2009). However, one should differentiate between the tools one uses to investigate the world and the world itself. While tools can change with advances in technology, the fact that scientists can identify patterns at particular scales (i.e., at some scales and not others) suggests that these patterns are not an artifact of the method of observation. If the patterns were indeed just an artifact of the method of identification, then scientists should be able to identify patterns at arbitrarily chosen scales. However, this is not the case: scientists can identify patterns only at specific scales, and this suggests that the world, rather than the scientists and the tools they use, imposes constraints on the types of patterns that can be identified.[27]

[27] Woodward also suggests that "true empirical assumptions will not license inconsistent phenomena claims from the same data" (Woodward 2011, 175). For example, in atmospheric dynamics, if data are collected and analyzed at a particular spatiotemporal scale (with a particular level of accuracy), then the confidence in the use of particular scale related assumptions will be reinforced if different research groups should identify the same regularities or irregularities in the data set at a given scale.

2.3.2 Circularity

The second potential problem for my account is that the convergence of the data-to-phenomena and theory-to-phenomena inferences for gaining confidence in the successful isolation of a phenomenon may be misleading because of a circularity: if theoretical elements are involved in the data-to-phenomena inference, and/or vice-versa, then the fact that the data-to-phenomena and the theory to phenomena inferences converge is no sign of confidence in the fact that scientists have identified a genuine phenomenon.[28]

In my view, the circularity is not vicious. This is because the scale assumptions involved in the data- and theory-to-phenomena inferences are motivated differently. In section 2.2.3, I have briefly described how dimensional analysis is used to obtain a simplified set of equations that describes large scale atmospheric motions. The empirical input into this derivation enters at the stage of evaluating the quantitative value for the length scale L and the velocity scale U of the phenomenon (the "order of magnitude"). The order of these parameters then allows for an estimate of the order of magnitude of the RN. To obtain the equations of motion for the large-scale system, several further steps need to be taken. Once the variables in the equation are written in terms of the non-dimensional parameters, one needs to choose particular ordering relations between the parameters before expanding the terms in the equation and discarding the terms that are not of the order of interest. An example of such an ordering relation is letting ε tend to zero while letting $\varepsilon r_0/L$ remain of the order of one. r_0 is the radius of the earth, and all other terms are defined as in section 2.2.3. This limiting ordering relation identifies large scale phenomena the vorticity of

[28] I understand that when theoretical elements occur in the data-to-phenomena inference and empirical elements in the theory-to-phenomena inference, it may become difficult to categorize them as two distinct types of inference. I will not address this worry in this dissertation.

which is driven both by planetary (driven by the rotation of the earth) and relative (driven by the local features of the phenomenon, in particular its size *L*) vorticity.

While the choice of ordering relation is determined by the observational evidence that determines the values for *L* and *U*, this does not suffice. As Pedlosky explains:

> We must, in addition, bring to the scaling analysis an intuitive appreciation that certain ordering relationships are highlighted by the way they distinguish important and natural physical balances that are physically relevant to the phenomena of interest. (Pedlosky 1987, 346)

These are balances such as the geostrophic balance, described in section 2.2.3. The ordering relation is ultimately at the heart of the theory to phenomena inference: it is based on the scale existence assumption, i.e., that phenomena occur at particular scales (see Pedlosky (1987, 340)), and it is the relation that allows to eliminate irrelevant terms from the equations of motion. The assumption about the scale of motion is thus part of a complex theoretical apparatus.

Here the circularity is not vicious. In the case of the empirical input in the theory-to-phenomena inference, the quantitative value of *L* and *U* is considered meaningful only if the ordering relation that is associated with them reflects a physically meaningful balance. As Pedlosky states in the above passage, different ordering relations highlight different processes in the atmosphere. On the other hand, the theoretical interpretation of the physically meaningful balances is fairly independent of the observed values of *L* and *U*, as the interpretation is guided by the geometrical boundaries described above. In the example I have given, the Coriolis force plays a dominant role in the balancing of forces, and the geostrophic balance is a balance that reflects this contribution in the dynamics of phenomena at the large scale.

In the case of the theoretical input in the data-to-phenomena inference, the physical interpretation of the ordering relation (i.e., the geostrophic balance) provides a theoretical justification for the data-to-phenomena inference: “[The] recognition of the particular meaning of this ordering relation then gives confidence that the observed numerical relations between parameters are not merely fortuitous” (Pedlosky 1987, 346). The numerical relations are obtained by plugging in values obtained by observations in the terms of the dimensionless parameters. This illustrates that the data-to-phenomena inference and the theory-to-phenomena inference are in a non-viciously circular feedback relation. The empirical data-to-phenomena inference, in this case, involves choosing the scale at which to average the spectrum for a particular parameter (e.g., the wind velocity) in order to obtain a characteristic scale. That value for the characteristic scale allows the theoretician to determine the order of the dimensionless parameters used to obtain the equations of motion for large scale phenomena. Together with some related physical interpretation of the system (e.g., geostrophic balance), the theoretician can obtain a particular (limiting) ordering relation between the parameters that allows her to obtain the new set of equations of motion. The physical interpretation of the ordering relation then gives confidence in the numerical values observed for the parameters at the scale of motion considered.

2.3.3 Empirical and Normative Assumptions

Peschard (2012a, 2012b) argues that important assumptions that identify the target system as such are not empirical, but normative. She calls these normative judgments “relevance judgments,” and define them as follows:

> Relevance judgments . . . are judgments of what quantity should be measured, or, correlatively, what effect should be taken into account by a measurement procedure. (Peschard 2012a, 750)

Relevance judgments are judgments about what relevant parameters and what effects of the parameters need to be taken into account to represent a target system. Peschard claims that relevance judgments are not empirical:

> But suppose now that the claim under test cannot be interpreted as being about a data-generating procedure. The issue then arises of what data-generating procedure is adequate for the test and generates data that qualify as benchmark for the evaluation of the claim. . . . The reliability of a given experimental procedure will depend on the answer to the question. This answer is a judgment that specifies what sort of data need to be acquired and will need to be accounted for by a model about the crime rate or the receptive field [the phenomenon]. It is not an empirical judgment; it is what we have called a 'relevance judgment. (Peschard 2012b, 758)

She distinguishes these types of judgments from claims about the reliability of data-gathering procedures, which, she argues, are empirical claims that can be tested only when phenomena are already established. Empirical assumptions are therefore assumptions that occur later in the modeling process, when the target system has already been identified.

If Peschard is right, then the scale existence and separation assumptions are not important for the process of identifying phenomena. However, my thesis has been that empirical assumptions are important for this process. The apparent conflict between these two views can be resolved if we compare the role of relevance judgments. As Peschard says, relevance judgments play at least two roles: selecting parameters that describe phenomena as well as the effects of the parameters, which also describe phenomena. The scale existence and separation assumptions play the same roles. Non-empirical, relevance judgments are therefore not the only types of judgments involved in the process of identifying phenomena.

2.4 THE PURPOSE OF IDEALIZATION

The scale existence and scale separation assumptions play an idealizing and abstracting role in the data-to-phenomena and in the theory-to-phenomena inferences. In this section I will argue that these idealizing assumptions are *essential* for the target system identification process.

Weisberg (2013) argues that there are many kinds of idealization, and different aims of the scientists may require different kinds of idealization. The three kinds of idealizations that Weisberg identifies include the ones mentioned above: Galilean Idealization, Minimalist Idealization and Multiple Model Idealization. He suggests that Galilean Idealizations (McMullin 1985) are pragmatically motivated idealizations, which are dependent on the status of the theoretical apparatus available to the scientist. Further, these idealizations are a distortion of the target system that may be removed as the science advances (Weisberg 2013, 99–100). Minimalist Idealizations, on the other hand, are idealizations that are part of models that include essential causal factors of the target system, and are in this sense ineliminable. Weisberg associates this view with Strevens (2004) and Batterman (2001, 2002, 2009). The two accounts of idealizations however are rather different, as for Strevens idealizations are non-difference making features of the model (Strevens 2004, 174), while for Batterman idealizations are procedures involved in obtaining a minimal model of a universality class that contains features that are common to several systems (Batterman 2002, 35). The common denominator of these two accounts is that the idealization is part of the explanatory text of the phenomena that the model represents. The third kind of idealization Weisberg introduces is the Multiple Model Idealization. In this case, idealizations are introduced to build inconsistent models of the same phenomenon. However, Weisberg argues, the inconsistency between models is warranted as the models aim to explain different aspects of the phenomenon of interest (Weisberg 2013, 103–105). While this taxonomy might be useful for descriptive purposes, it does not shed light on how the idealization process isolates the target system from its environment.

Idealizations are in fact usually discussed in the context of already established target systems: once the target system is established, an idealization is introduced to eliminate intractable aspects of the target system or to highlight particular aspects thereof. However, in the previous sections I have shown that idealizations feature importantly in the target system identification process: they are responsible for identifying relevant parameters that describe the target system. Thus, these idealizing assumptions are necessary for obtaining knowledge about the world.

2.5 CONCLUSION

At the beginning of this chapter, I have asked what kind of assumptions are involved in the data-to-phenomena and theory-to-phenomena inferences, and how these assumptions contribute to the practice of identifying phenomena. In describing the idealizing role of scale existence and scale separation assumptions, I have argued that they are necessary empirical assumptions in the target identification process: these assumptions are one of the building blocks of the data-to-phenomena and theory-to-phenomena inferences. They distinguish signal from noise in the first case, and guide scale analysis, and consequently the (formal) choice of relevant parameters in the second case. This suggests, in line with general features of Batterman’s view, that idealizations are part of a process that isolates stable patterns in the world. Idealizations and abstractions are ineliminable in so far as they are the isolate the relevant aspects of the (parameters that describe the) target system.[29] In particular, spatiotemporal idealizations are fundamental to the modeling process in because they isolate the target system from its environment.[30] Under this reading of the science, models can represent and explain the world in virtue of idealizations, not in spite of them.

This account of the role of idealization in the data-to-phenomena and theory-to-phenomena inference illustrates that discussing the relation between data, theory and phenomena can provide insight into the role of idealizations in modeling. More precisely, the analysis of the structure of the inferential relation from data and from theory to phenomena illustrates how idealizations feature in the process of identifying target systems, and, as a consequence, how models of target systems can provide reliable knowledge of the world.

[29] The difference between my account and Batterman's is that mine is slightly more general and is not tied to a philosophical account of explanation.
[30] In the context of Peschard's work, they provide the empirical basis for the isolation of the target system from its environment.

3.0 DATA AND THEORY TO PHENOMENA INFERENCES

3.1 INTRODUCTION

In the previous chapter, I have argued that assumptions about the scale of phenomena are empirical assumptions that play an important role in the identification of targets. These assumptions are idealizing and abstracting and can occur both in theory-to-phenomena and data-to-phenomena inferences. This chapter further illustrates these claims. The last chapter analyzed the assumptions for a very well-established case. The identification of large scale phenomena such as Rossby waves and the derivation of the modeling equations for these phenomena are included in most fluid dynamics textbooks. In this chapter, I will take an historical approach and discuss the development of the characterization of climate phenomena over time. As in chapter 2, I will analyze *the scale existence assumption* and *the scale separation assumption*. The latter relies on the former; one needs to first assume that a phenomenon exists at a characteristic scale before establishing a criterion for differentiating one phenomenon from the other.

The data-to-phenomena inference is studied for the El Niño Southern Oscillation (ENSO) phenomenon, where the spatial arrangement of the observation network and temporal sampling of data play a fundamental role in understanding ENSO phenomenon and identifying its main components. I will illustrate two landmark steps that improved the scientists' understanding of ENSO.

In the fifties, a prominent meteorologist Bjerknes suggested a general mechanism for ENSO. While this was a crucial step towards identifying ENSO as the phenomenon that we know today, it was not until an array of buoys that was specifically aimed at resolving particular scales that a more satisfactory characterization of ENSO arose. Here is the second step: a failure to detect the onset of a particularly strong ENSO event in the early eighties prompted the development of a research program that aimed at creating an observation network that would resolve the scales at which the main components of ENSO were thought to take place. Important methodological steps of this program involved paying attention to the role of spatiotemporal scales in data collection and data analysis for the identification of phenomena.

The need for an array of observation stations that takes particular scales into account suggests that the data to phenomena inference involves nontrivial assumptions about the scales at which phenomena and its components are identified. While Bjerknes had already identified ENSO as a phenomenon occurring on the spatial scale of the Pacific Ocean and peaking during January, he could not provide the scientific community with the conceptual tools to identify the occurrence of all subsequent ENSO events (most notably the ENSO of 1982–83). For a clear signal from ENSO to be detected, an organized data collection array was needed that would provide data sets at spatial and temporal scales fine grained enough to distinguish the signal of the phenomenon and its components from the noise coming from other scales. This process illustrates that scale related assumptions are pervasive in the process of identifying a phenomenon and its' component.

The theory-to-phenomena inference is studied for the case of the parametrization scheme introduced by Arakawa and Schubert in 1974. This parametrization scheme is used to parametrize cloud physics that occurs at scales smaller than cumulus clouds of a larger system. An example of a larger system is a hurricane. A core assumption in the justification of the Arakawa-Schubert parametrization scheme is the scale separation assumption: the parametrization is considered valid because the cloud physics occurs at scales much smaller than the scale of interaction of collections of cumulus clouds within a hurricane, or another large scale perturbation. This assumption is essential not only for obtaining models that predict the behavior of phenomena in which the parametrization is employed, but it is also used as a crucial tool for understanding the dynamics of such phenomena. This parametrization scheme is an example of how the scale separation assumption is used to justify what physics gets included into a model of a phenomenon (interaction between ensembles of cumulus clouds and the other components of the phenomenon) and which is not (physics of single cumulus clouds).

3.2 ENSO: THE ASSUMPTIONS IN THE DATA TO PHENOMENA INFERENCE

The definition of ENSO that is most widely accepted by today's scientific community is the following:

> El Niño can be said to occur if 5-month running means of sea surface temperature (SST) anomalies in the Nino 3.4 region (5N-5S, 120-170W) exceed 0.4C for 6 months or more. (Trenberth 1997, 2771)

This definition is based on the assumption that a time averaged mean in a particular region is characteristic of the phenomenon, which spans over the whole Pacific over a period of more than 5 months, if one considers its onset, its peak and its dissolution. The question that this section will explore is what role the scale existence assumption plays in the driving the data collection and analysis that lead to this definition. In other words, what role does the scale existence assumption play in data to phenomena inference that identifies a phenomenon defined by the above quote?

I will start answering this question by illustrating theoretical framework that lead to a drastic change in the way in which oceanic data was collected in order to identify the main components of ENSO. This is the framework provided by the important contributions of Bjerknes (1969). I will show how his own contribution also rests on the scale existence assumption. I will then describe the new data collection experiment that followed the realization that that framework was not satisfactory–it did not allow scientists to recognize the onset of a particularly significant ENSO event in the early eighties. Again, scale related assumptions played a major role in this process.

3.2.1 Bjerknes

Bjerknes is recognized as the first meteorologist to define ENSO as a basin wide phenomenon that involves the interaction between atmospheric and oceanic components. Bjerknes identified a recurring anomaly from average pressure and sea surface temperature values of the Pacific Ocean in particular seasons across different years, and this anomaly occurred simultaneously for atmospheric and oceanic parameters. He associated these anomalies with a disruption of a circulation of air above the Pacific called the "Walker Circulation."[31] The anomalies are associated with above average rainfall and a characteristic low-pressure pattern across the pacific basin for the month of January (basin-wide spatial scale, one month long temporal scale).

The scale related assumption at the basis of his description of ENSO is the scale existence assumption, namely that there is a characteristic scale at which the phenomenon can be identified.

[31] For a historical and philosophical discussion of the discovery of the Walker Circulation, see Katz (2002), Pincock (2009).

The scale existence assumption is justified by the fact that the anomalies observed in the mean field at the scale that defines the phenomenon are recurrent: there are semi regular intervals at which the variability that defines ENSO occurs. In particular, ENSO anomalies as defined by Bjerknes recur in the month of January (though not necessarily at regular intervals).

The assumption about the existence of a scale at which the phenomenon occurs appears in Bjerknes' analysis of the air pressure, rainfall and sea temperature data of the years 1956–1967 for stations in the Pacific basin area:

> *The selection of the January maps of 1963, 1964, 1965, 1966, and 1967 brings into focus the anomalies of the large-scale flow* that evolved together with the extreme coolness and dryness of the central Pacific equatorial belt in January 1963, 1965, and 1967 and the extreme warmth and rain surplus of the same belt in January 1964 and 1966. (Bjerknes 1969, 164; my emphasis)

Bjerknes states that he is explicitly choosing maps that represent the large scale flow at specific spatial and temporal intervals. This choice of maps allows him to show that the winters of 1963–64 and the winter of 1965–66 are associated with an anomaly in the flow pattern spanning the whole pacific basin that occurs at the same time as what was previously thought to be ENSO events.[32] The anomaly that emerges from the analysis of the data for these years, he concludes, is a disturbance large enough to provide a persistent disruption of the large scale pressure fields and eliminating the typical wind patterns (Bjerknes 1969, 165).

[32] ENSO events were previously associated with a change in rain patterns along the Western coast of the Americas.

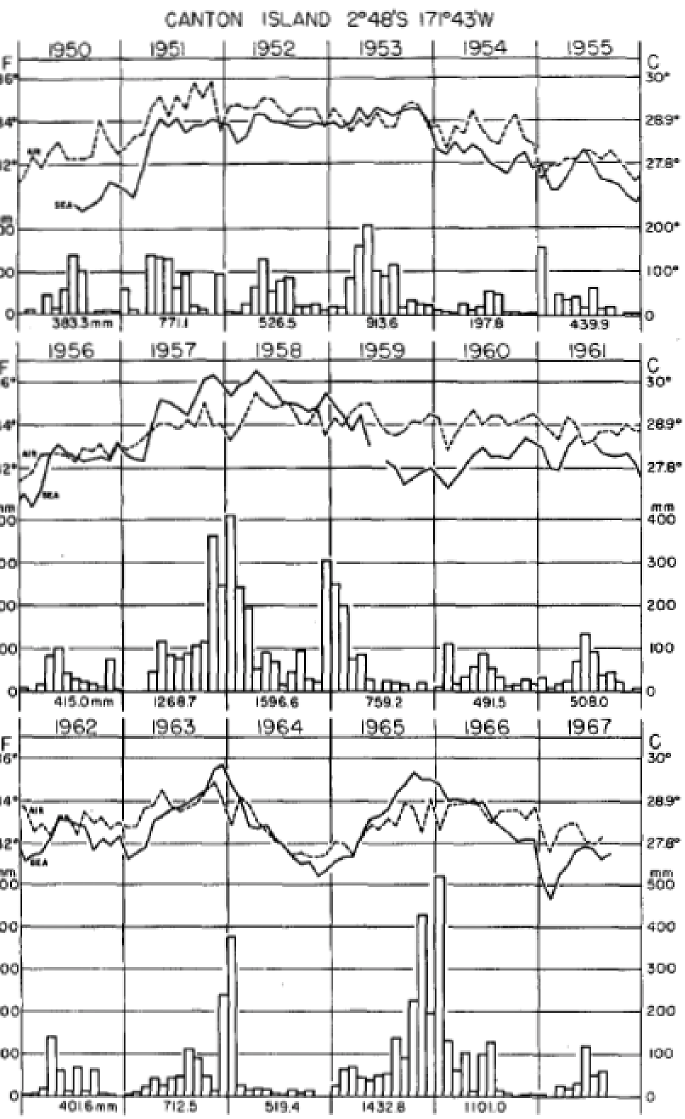

Figure 3.1. Time series of monthly air and sea temperature and of monthly precipitation at Canton Island from 1950 to 1967. From Bjerknes (1969).

The figures provided illustrate the anomaly identified by Bjerknes. Figure 3.1 shows the monthly average values for temperature and rainfall in Canton (Canton Island is an Island of the Kiribati Archipelago, in the mid-west equatorial Pacific). There are clear peaks in the average values for December 1957 to January 1958, and in the average values for the December–January period in the years 1963-64 and 1965-66. The important feature of these data series is that they reveal peaks in rainfall and temperature that occur more or less simultaneously, always in the same season. Further, these peaks occur in the same season as El Nino events registered on the Western Coast of South America, i.e. a heavy rainfall along the northern coast of western South America and a warming of the sea surface temperature (SST) along that coast. The resulting correlation between different locations in rainfall and temperature increase indicates that the anomaly is associated with a phenomenon that occurs at the scale of the Pacific Ocean. To corroborate this claim–i.e. that the peaks in rainfall are part of one phenomenon and not just two phenomena occurring simultaneously in different regions of the globe–maps of the whole Pacific basin are produced.

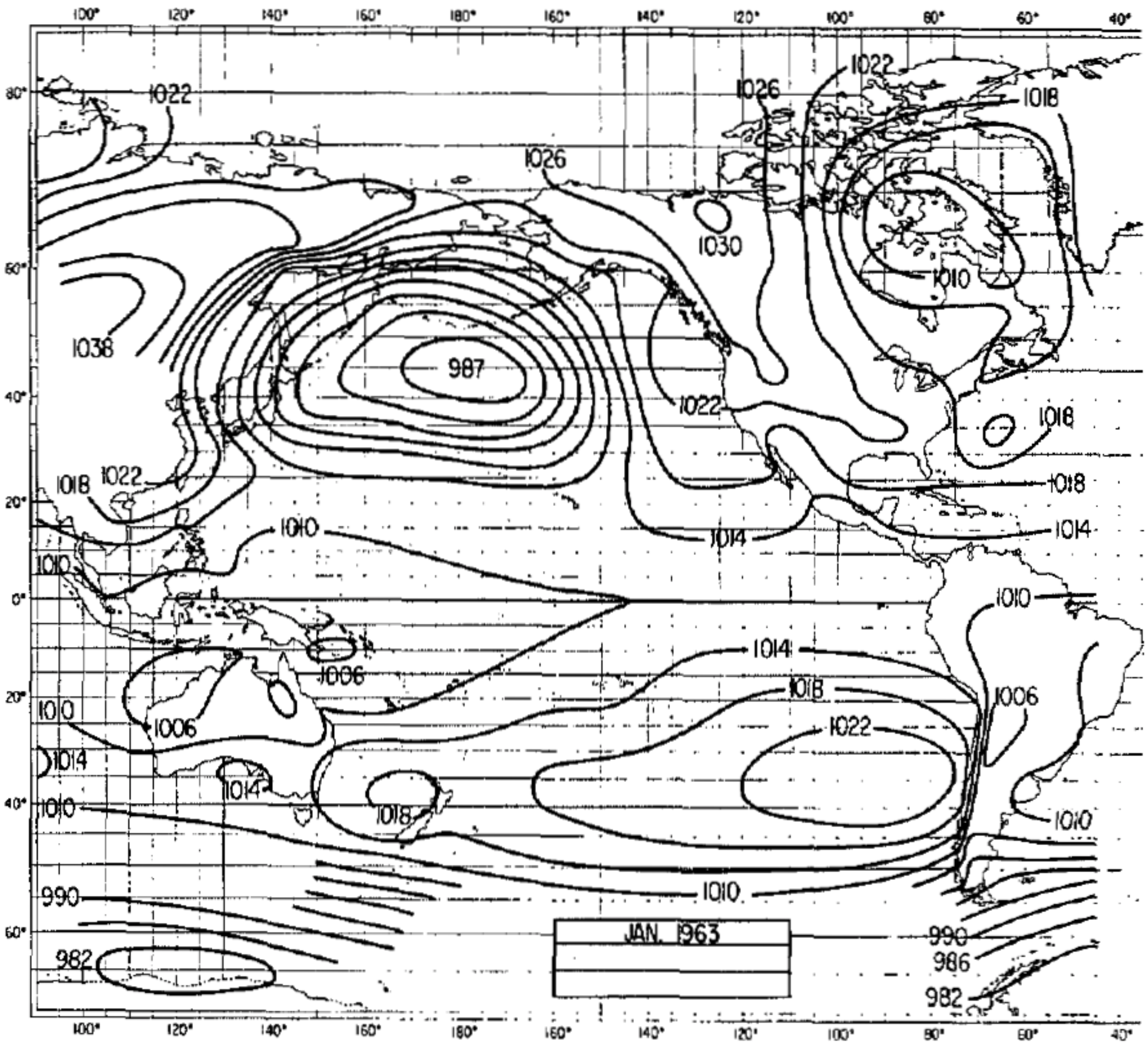

Figure 3.2. January 1963 distribution of pressure (millibars) at sea level. From Bjerknes (1969).

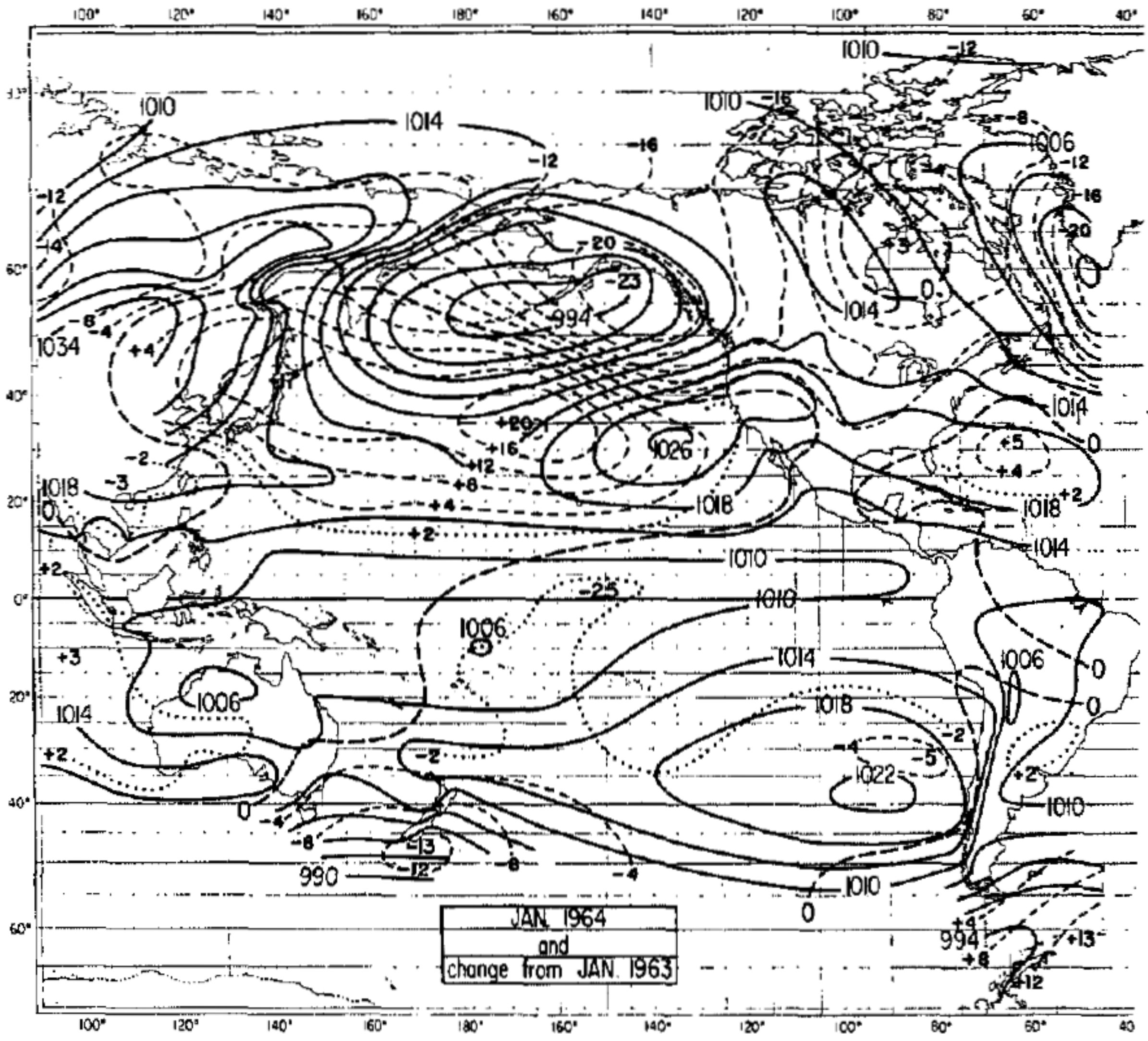

Figure 3.3. January 1964 distribution of pressure (millibars) at sea level. Change from January 1963 in dashed isallobars. From Bjerknes (1969).

Figures 3.2 and 3.3 show average pressure patterns for January for the years 1963– for which there is no ENSO, dry and cool Eastern Pacific conditions–and 1964–for which there is an ENSO event, hot and wet Eastern Pacific conditions–at the scale of interest. The comparison between the maps that resolve this particular scale reveal a clear shift in the pressure pattern along the equator in the

Pacific basin: the 1010mb isallobar,[33] that during January 1963 only extends eastward a third of the way along the equator, extends all the way to the coast of Colombia and Peru during the January of 1964, bringing heat and humidity to an otherwise cool and dry region. A similar change in pattern can be observed for the same region in the Januaries of 1965 and 1966 (Bjerknes 1969, 164).

The scale existence assumption is evident in Bjerknes' use of the data series and the maps to identify ENSO as a basin-wide phenomenon rather than a phenomenon that only occurs on the Western coast of the Americas: The maps of the observed parameters at a particular scale, and in particular monthly averages, show that there is a recurring anomaly of the pressure and temperature fields, and average rainfall amount resolved by this particular choice of average. The recurrence of this anomaly prompts Bjerknes to claim that there is a stable phenomenon occurring quasi periodically over the same region at a particular time scale. Fluctuations in physical quantities of interest can occur on larger and smaller scales, but at the monthly spatiotemporal scale these variations can be given physical meaning.

The identification of the scale at which ENSO occurs, however, does not suffice to fully characterize the phenomenon. Bjerknes himself states that to identify other components of the phenomenon, data should also be analyzed and collected at smaller scales–which implies that components of the phenomenon can and do occur at other scales:

> The time lag of the large-scale atmospheric response to the initial anomaly of heat input from the equatorial ocean appears to be quite small and would have to be identified by daily instead of monthly basic data. (Bjerknes 1969, 165)

[33] An isallobar is a line on a map connecting points at which the atmospheric pressure has changed by an equal amount during a specified interval of time.

Bjerknes is stating that the anomaly in the oceanic conditions starts slightly earlier than the anomaly in atmospheric conditions, suggesting that there must be some processes that occur at smaller scales that connect the two anomalies. Bjerknes suggests that the investigation of how anomalies in the pressure and sea surface temperature fields are related should be conducted at a scale smaller than the one he has been investigating.[34]

There is a difference in the scale of resolution depending on what is being sought about the target system. To identify the extent of ENSO, it was sufficient to identify a correlation between the large scale anomalies in the relevant physical quantities: a recurrent pattern was observed at the scale of the pacific basin both in pressure and sea surface temperature values. However, to fully appreciate the extent of the phenomenon and the scale at which it occurs, a different temporal scale needs to be investigated. In the next section I will illustrate how a different strategy for data collection needed to be developed in order to be able to detect ENSO before it reaches its peak.

The importance of the scale existence assumption in this process is that the scale of the phenomenon as a whole is the starting point for identifying the phenomenon and starting to build a model thereof. The data gets collected and analyzed at the scale at which the phenomenon is assumed to occur, and a similar procedure is implemented when components of a phenomenon are identified—if they occur at scales different from the scale of the phenomenon as a whole, which occurs often in climate modeling. I will now describe how paying attention to the scale of the phenomenon and its component lead to the development of an observation network that lead to a more effective characterization of ENSO.

[34] The details of the mechanism that connects the oceanic and atmospheric components of ENSO are still disputed today (Guilyardi et al. 2009).

3.2.2 TAO and TOGA

The need for a principled observation network that would resolve different scales of motion arose not long after the recognition that ENSO is a coupled ocean-atmosphere phenomenon–a phenomenon that has both atmospheric and oceanic components. In the years 1982–1983, the scientific community realized that the models that were used to predict ENSO failed to capture some of the salient components of the phenomenon. That year's ENSO "was neither predicted nor detected until nearly at its peak" (McPhaden 2006, 85). The lack of being able to predict the phenomenon indicated to the scientific community interested in studying ENSO that more data needed to be collected, and given the complexity of the phenomenon, data collection should be planned carefully and systematically. In McPhaden's own words recollecting the status of ENSO research at the time—he served as of the lead oceanographers of the Tropical Ocean Global Atmosphere (TOGA) experiment:

> The sparsity of accurate oceanic measurements in 1982 exposed not only our ignorance about El Niño's complexity but also *the gross inadequacy of existing observing systems to measure and describe it.* (McPhaden 2006, 86; my emphasis)

There were only few stationary buoys that could collect time series data of sea surface temperature and most data was collected by commercial ships and the newly developed satellites (McPhaden 2006, 85). In order to detect an ENSO event, warmer than average sea surface temperature should be observed for a period of time. This, however, did not occur. Satellite data for 1982 year showed a cold temperature bias caused by the eruption of a volcano in southern Mexico, which was mistaken for cloud cover. The warmer temperature values of the data collected in situ was discarded as being outliers (McPhaden 2006, 85), which lead scientists to conclude that that year normal conditions would prevail, and no ENSO event would occur. However, after a couple of months, they were proven wrong when they observed the peak of an ENSO event.

The satellite data had proven unreliable–the source of the bias was not correctly attributed to its cause until later. Scientists had been relying too heavily on the satellite data: the data coming from the few stationary buoys that were actually indicating the warmer conditions, which should have warned the scientists of the onset of an ENSO event, was discarded as an outlier. The inadequacy of the data collection and analysis that year indicated that in order to better understand ENSO, more stationary, long term, observation stations were needed across the Pacific basin. The goal was to build an observation network that could clearly distinguish the signal coming from ENSO from signals coming from other phenomena at different scales.

The observation system implemented for the purpose to better identify ENSO and its components throughout the 80s and 90s was the TOGA program. Its wider scope was to support climate studies from the seasonal to the interannual timescale across the Pacific basin (McPhaden et al. 1998, 169). The program brought two main scale-related innovations with respect to past data collection. First, it would be a long term program (of at least 10 years). This time scale was thought to be the scale at which physically significant interannual oscillations recur at the spatial scale of the Pacific basin. Second, it would have a fixed network of observation stations (instead of drifters—moving data-collecting buoys—or data collected by ships). A fixed network would provide a time-dependent picture of the basin scale variations of parameters of interest:

> This observing system was to provide data on a basin scale for at least 10 years without significant temporal gaps, so that a continuous record of climate variability could be assembled. Ten years was considered the minimum length of time needed

> for a comprehensive study of interannual variability, the dominant mode of which was the ENSO cycle. (McPhaden et al. 1998, 171)

Here, McPhaden is highlighting the importance of an organized array of observations aimed at detecting physically meaningful scales of motion, i.e. fluctuations of the wind, pressure and temperature fields associated with phenomena of interest such as ENSO.

This reasoning is in line with Stommel's (1963) observation described in chapter 1, according to which different phenomena occur at different scales, and a systematic observation network can provide data that can be analyzed more easily. Further, a systematic array that can collect data at many time scales at a fixed spatial scale can help reduce difficulties tied to scale detection: in order to resolve peaks in the data that are associated with phenomena of interest, the noise needs to be resolved in order to avoid contamination and, as a consequence, misguided claims about the phenomenon of interest.

> It was recognized at the start of TOGA that although ENSO is predominantly a large-scale, interannual perturbation of the climate system, it could not be effectively observed without taking into account smaller-scale, higher frequency fluctuations. There is a broad spectrum of variability in both the ocean and the atmosphere that represents a broad spectrum of geophysical noise in estimates of climate signals. *Noise contamination can arise because of inadequate sampling in space and/or time* which will alias energy from high-frequency, small-scale fluctuations into the lower frequencies and larger scales of climatic interest. (McPhaden et al. 1998, 174; my emphasis)

Signals from phenomena occurring at other scales can muddy the signal from the ENSO phenomenon, and in order to detect the ENSO signal, this noise needs to be correctly identified.

If the signal of a phenomenon can be identified at a particular scale (the scale existence assumption), the associated dynamical field can be characterized in the following way:

$$\Psi(\boldsymbol{x})_{obs} = \Psi(\boldsymbol{x}) + \varepsilon_{\Psi(\boldsymbol{x})}$$

where the left-hand side of the equation represents the observed field (i.e. the time series of a parameter **x** = (x, y, z, t) in space and time) and the right hand side represents the separation in scale of the field of the phenomenon itself $\Psi(\boldsymbol{x})$ and the noise on the field coming from measurement error and variability at unresolved scales $\varepsilon_{\Psi(\boldsymbol{x})}$. Unless $\varepsilon_{\Psi(\boldsymbol{x})}$ is randomly distributed, the operation of separating signal and noise is nontrivial. If multiple scales can be resolved, and there are two components at two different scales, then the observed field can be characterized in the following way:[35]

$$\Psi(\boldsymbol{x})_{obs} = \Psi(\boldsymbol{x}) + \Psi'(\boldsymbol{x}) + \varepsilon_{\Psi(\boldsymbol{x})}$$

Where $\Psi'(\boldsymbol{x})$ is the smaller scale signal, and the rest is defined as above. The smaller scale signal is the anomaly on the larger scale field, and, like measurement error and unresolved scales, can introduce error in the identification of the large scale signal.

An instance of the of identifying a signal of a phenomenon from a data series is the data on which Bjerknes relies to identify ENSO. As described above, this data can sometimes be misleading:

> Various versions of the SOI [Southern Oscillation Index] exist although, in recent years, most deal only with atmospheric pressures and usually only those of Darwin and Tahiti. In using the SOI based on just two stations, it must be recognized that there are many small scale and high frequency phenomena in the atmosphere, such as the Madden Julian Oscillation, that can influence the pressures at stations

[35] See Gray and Riser (2015, 4341).

involved in forming the SOI but that do not reflect the SOI itself. (Trenberth 1997, 2773)

The Southern Oscillation Index is an index for the state of pressure and sea surface temperature fields in the Pacific and Indian oceans. Particular values for the SOI would indicate a change in the Southern Oscillations, and as a consequence the occurrence of ENSO. Trenberth is stating that a SOI only based on a time series coming only from two spatial locations (at the western boundary––Darwin—and in the center—Tahiti—) could not provide reliable values for ENSO due to contamination from noise at other scales.

To have a better index for ENSO, enough data needs to be systematically collected in order to separate the genuine phenomenon from noise or signals at other scales. Table 3.1 is a chart of the scales of observations developed by scientists at TOGA to obtain reliable monthly averages of the parameters of interest.

Table 3.1. TOGA data requirements. From McPhaden et al. (1998). Reprinted with permission from AGU.

Parameter	Horizontal (Vertical) Resolution	Time Resolution, days	Accuracy
Upper air winds	500 km (two levels: 900 and 200 mbar)	1	3 m s^{-1}
Tropical wind profiles	2500 km (100 mbar)	1	3 m s^{-1}
Surface pressure	1200 km	1	1 mb
Total-column precipitable water	500 km	1	0.5 g/cm^2
Area-averaged total precipitation	2° latitude × 10° longitude	5	1 cm
Global sea surface temperature	2° latitude × 2° longitude	30	0.5°K
Tropical sea surface temperature	1° latitude × 1° longitude	15	0.3°–0.5°K
Tropical surface wind[a]	2° latitude × 10° longitude	30	0.5 m s^{-1}
Tropical surface wind stress[a]	2° latitude × 10° longitude	30	0.01 Pa
Surface net radiation	2° latitude × 10° longitude	30	10 W m^{-2}
Surface humidity	2° latitude × 10° longitude	30	0.5 g kg^{-1}
Surface air temperature	2° latitude × 10° longitude	30	0.5 K
Tropical sea level	as permitted[b]	1	2 cm
Tropical ocean subsurface temperature and salinity	as permitted[c]	as permitted[c]	as permitted[c]
Tropical ocean surface salinity	2° latitude × 10° longitude	30	0.03 PSU
Tropical ocean-surface circulation	2° latitude × 10° longitude	30	0.1 m s^{-1}
Subsurface equatorial currents	30° longitude (five levels)	as recorded	0.1 m s^{-1}

[a]While the accuracy requirements given are for 30-day averages, daily values are required for resolution of 30- to 60-day oscillations.
[b]As permitted by the existence of suitable sites and satellite altimetry.
[c]As permitted by appropriate in situ measurements techniques.

While this arrangement changed and still changes today, it illustrates the importance of the assumption that phenomena (and other components) occur at characteristic spatiotemporal scales and the individuation of a signal at a particular scale is a crucial step for identifying phenomena from data. The assumption does in fact lie at the basis of how we choose to lay out an observation network.

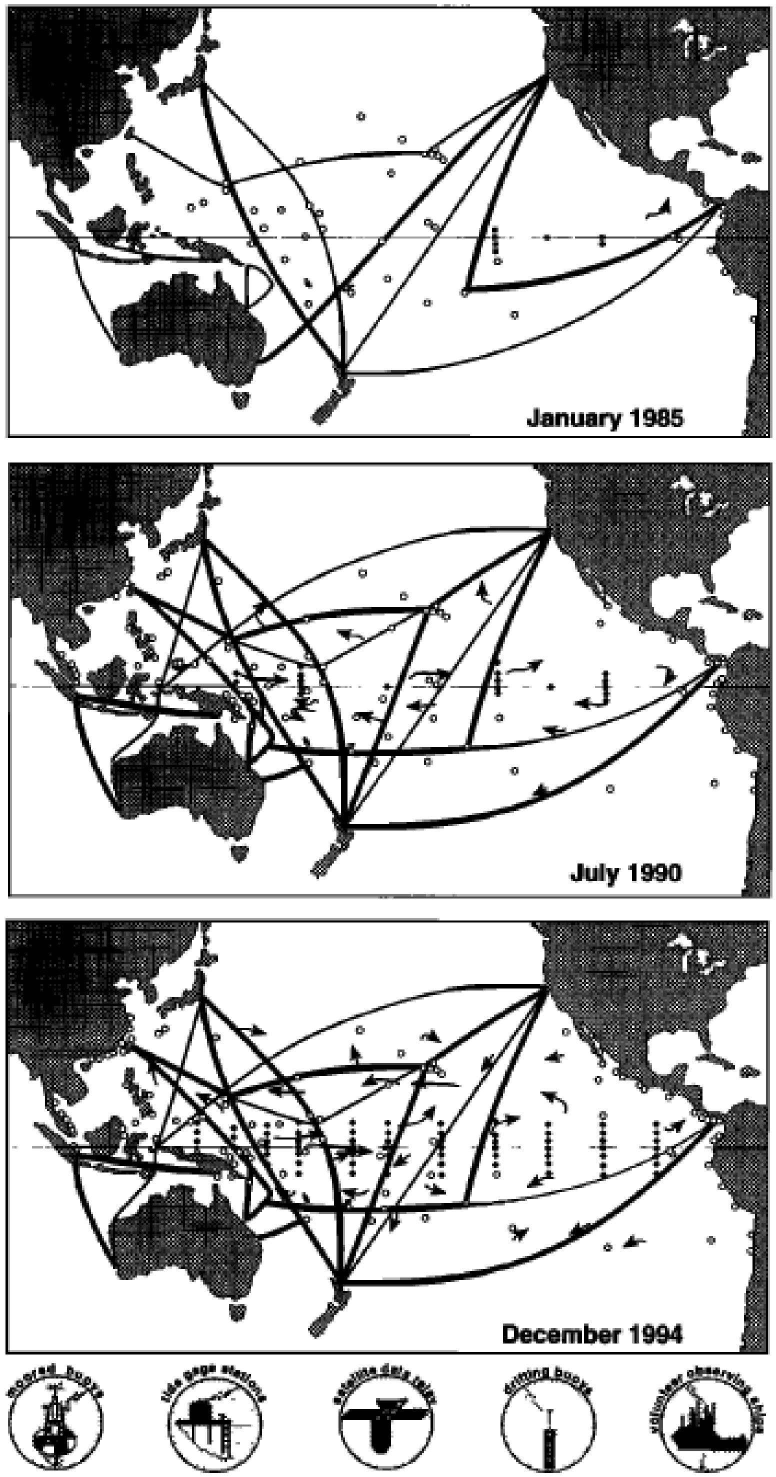

Figure 3.4. TOGA in situ ocean observing system. From McPhaden et al. (1998). Reprinted with permission from AGU.

The TOGA data collection experiment and its analysis lead to important developments in the characterization of ENSO, which eventually lead to the definition of ENSO given at the beginning of this section. In their 1997 research note, Barnston et al. suggest that regions previously thought to be relevant for identifying early components of ENSO might have been chosen only because of their convenience for data collection (Barnston et al. 1997, 368).[36] However, new data collected in the years of implementation of the TOGA experiment and data collected subsequently by the array allowed for a clearer analysis and a new and better outlook on the dynamical fields associated with ENSO (Barnston et al. 1997, 371). For example, a region in the Pacific labeled as "region 4" did not show a clear signal because it is a region in which there is a high degree of noise and low degree of variability (Barnston et al. 1997, 377). This means that the scale of the phenomenon was not clearly identifiable in that region when scarce data was available.

The role of the scale existence assumption in the data to phenomena inference thus is the following: it serves at as a guide for scientists to identify the main components of the system, in order to obtain a satisfactory characterization of the phenomenon. In the case of ENSO, this assumption is at the basis of Bjerknes' identification of the correlation between atmospheric and oceanic components of ENSO, which lead to various physical descriptions of the dynamics of ENSO. Later, the identification of a dominant scale of variability lead to the identification of critical areas of maximum variability for parameters associated with ENSO, which today defines the occurrence of an ENSO event. The TOGA experiment was instrumental: it provided a solid empirical basis for individuating the scale at which ENSO occurs.

[36] See also McPhaden et al. (1998). For a general overview on ENSO definitions, see Trenberth (1997).

3.3 ARAKAWA AND SCHUBERT: THE ASSUMPTIONS IN THE THEORY TO PHENOMENA INFERENCE

The scale existence and scale separation assumptions also feature prominently in the theory-to-phenomena inference. In this section I will focus on the scale separation assumption. This assumption relies on the scale existence assumption in so far as a phenomenon has to exist at a characteristic scale in order to observe a separation for a phenomenon at another scale.

In atmospheric dynamics, the scale separation assumption usually separates two components of a dynamical system, identifying principal and secondary ones. The secondary component is considered a disturbance on the large scale component, and that these disturbances occur on a scale much smaller than the larger scale. The disturbances that occur on the smaller scale can be parametrized, i.e. the details of their dynamics can effectively be suppressed in the large scale model. This allows scientists to isolate certain principal (relevant) components of the large scale model from smaller scale ones that are not directly relevant for describing its dynamics.[37] I will now analyze the role of the scale separation assumption in the parametrization of the dynamical details of organized cumulus convection in large scale phenomena, such as hurricanes and squall lines.[38] I will give a technical summary of the use and role of the scale separation assumption, while the philosophical significance of this parametrization scheme will be analyzed in more detail in the next chapter.

The most influential cloud parametrization scheme was introduced by Arakawa and Schubert (1974). While this parametrization scheme is used in numerical rather than analytic models of the atmosphere, it should still be regarded as an instance of the theory to phenomena

[37] I will expand on the epistemic significance of parametrization for modeling in the next chapter.
[38] A squall line is a band of storms and winds associated with cold fronts.

inference, as its derivation is mostly theoretical: Arakawa and Schubert explicitly state that their effort is directed at providing a theoretical framework for a parametrization that had previously been obtained with "a high degree of empiricism and intuition" (Arakawa and Schubert 1974, 674).

Organized cumulus convection is a mesoscale process (it can be observed at the mesoscale) and occurs when a number of cumulus clouds are present within a larger scale system. A central aspect of its modeling is how to account for clouds and their cumulative behavior. Should each cloud be modeled individually, or can the collective behavior of the clouds be parametrized in terms of variables of the large scale system? In this section I will show that there is not a simple or unique answer to this question.

While it is recognized that processes occur in clouds at many different scales, and that it is important to eventually understand and model all such processes, the parametrization is not merely sought for mathematical convenience, but has specific epistemic motivations. These motivations are in part to isolate and understand the recurrent behavior of organized convection at the scale at which it occurs. As a matter of fact, Arakawa and Schubert claim that modeling each singular cumulus cloud, a scale smaller than the one of organized cumulus convection, would be counterproductive for gaining knowledge about the target system:

> It should be emphasized here that the need for parametrizations is not limited to "numerical" models. Formulating the statistical behavior of small sale processes is needed for understanding large scale phenomena regardless of whether we use numerical, theoretical, or conceptual models. *Even under a hypothetical situation in which we have a model that resolves all scales, it alone does not automatically give us an understanding of scale interactions.* Understanding inevitably requires simplifications, including various levels of "parametrizations", either explicitly or

implicitly, which are quantitative statements on the statistical behavior of the processes involved. Parametrizations thus have their own scientific merits. (Arakawa 2004, 2496; my emphasis)

Parametrizations suppress degrees of freedom that are irrelevant for the purpose of isolating and understanding the behavior of the system of interest. Parametrizations also allow scientists to isolate phenomena and their components in order to study the way phenomena interact with other phenomena at different scales. The suppression of these degrees of freedom, therefore, is not only necessary because of computational constraints. As Arakawa points out, a model that resolves all scales, including the one at which cloud microphysics becomes relevant, does not necessarily provide a better model, i.e. a model that satisfactorily represents phenomena of interest or that provides reliable quantitative predictions.

Isolating components at their characteristic scales does two things: it identifies relevant components of the phenomenon of interest and it contributes to successful modeling of the interaction across components at different scales. One needs to be able to isolate components at particular scales to model how these components interact. The isolation of components, according to Arakawa, occurs by individuating the statistical behavior of small scale processes.

The scale separation assumption can be broken down into two components, one purely related to spatial scale of the system of interest, and the other related to the temporal scale of the system of interest. Both are presented in the argument found in Arakawa and Shubert’s 1974 original contribution. The first component is the spatial averaging assumption and provides the spatial scale (the size of a unit horizontal area) for the averaging of the collective behavior of cumulus clouds. The scale suggested by Arakawa and Schubert has to “be large enough to contain an ensemble of cumulus clouds but small enough to cover only a fraction of a large-scale disturbance” (Arakawa and Schubert 1974, 675). This means that the ensemble of cumulus clouds has to cover an area of the large scale system that is small relative to the system itself, but this area still has to contain multiple cumulus clouds. Further, within this area, the total area covered by a horizontal cross section of cumulus clouds cover must be small, i.e. within the cloud ensemble, the relative area occupied by clouds (versus cloud-free area) also has to be small (Arakawa and Schubert 1974, 677). This assumption somewhat constraints the applicability of the parametrization: the organized cumulus convection has to be dense enough (or not too dense) in order for this assumption to be satisfied. The area obtained by this assumption is what is sometimes called in the scientific and philosophical literature “Representative Element Volume” (see Batterman (2013, 262) and references therein), and getting the scale of this volume right is crucial for the physical argument that describes the process at that scale.

The time scale related assumption is that the time scale at which the convective processes of the cumulus clouds adjust to changes in the large scale environment (the phenomenon of interest) is sufficiently shorter than the time scale of the phenomenon (Arakawa and Schubert 1974, 691).

The spatial averaging assumption and the time scale separation assumption are at the basis of a physical quasi-equilibrium balance argument. This is a balance of the convective processes of the cloud with the large scale constraints posed on the convection itself (Arakawa and Schubert, 1974, 675, 691; Yano and Plant 2012, 7). At an appropriate scale, the cloud physics is in equilibrium with its environment (the large scale system), and, as a consequence, details of the cloud physics can be ignored. This argument is very similar to the notion of thermodynamic balance (Yano and Plant 2012, 6) and the continuum hypothesis in fluid dynamics (Emanuel 2007).

The (somewhat more) detailed argument is the following. A convective process involves the adiabatic lifting and sinking of a fluid element. The two parameters that describe this process are the vertical momentum $\rho\, w_c$ (where ρ is the fluid density w_c is the vertical velocity) and the buoyancy b. This process occurs only for a fraction of the total area of the phenomenon, namely the fraction covered by cumulus clouds σ_c. The spatial averaging assumption is applicable if $\sigma_c \ll 1$ which means that the area covered by cumulus clouds must be smaller than the total area taken into consideration. The total convective kinetic energy that is generated by the cumulus cloud ensemble is quantified by

$$A = \int_{z_B}^{z_T} \frac{\sigma_c\, \rho\, w_c}{\widehat{M}}\, b\, dz$$

where A is called the "cloud work function" (Arakawa and Schubert 1974, 687), z_B and z_T the cloud base and top respectively, and $\widehat{M}$ is a normalization factor, the rate of vertical mass transport of the convective process. The rate of change of kinetic energy of an ensemble of clouds is given by the following general equation

$$\frac{d}{dt} A_\lambda = F_{L,\lambda} - D_{c,\lambda}$$

Where λ accounts for the cloud type, $F_{L,\lambda}$ is the rate at which A_λ is generated by large scale process (the environment/large scale phenomenon) and $D_{c,\lambda}$is the rate at which A_λ is consumed by the convective process, i.e. the kinetic energy used up by the cloud during convection (see Yano and Plant 2012, 7).

The kinetic energy consumption rate of the convective process is given by

$$D_{c,\lambda} = \sum_{\lambda\prime} K_{\lambda\lambda\prime} \widehat{M_{\lambda\prime}}$$

where $K_{\lambda\lambda'}$ is the rate at which each unit of cloud base mass flux for the cloud type λ' contributes to the reduction of A_λ.[39] The convective quasi equilibrium assumption, a crucial assumption that provides the closure for Arakawa and Schubert's parametrization scheme, is

$$F_{L,\lambda} - D_{c,\lambda} = 0$$

which means that the cloud is in equilibrium with its environment: $\frac{d}{dt}A_\lambda = 0$. This approximation can be applied only if the time scale of the adjustment processes of the cloud τ_{ADJ} is much smaller than the time scale of the large scale process τ_{LS}. This is the time scale separation assumption mentioned above, and according to Arakawa and Schubert, it is "an assumption on parametrizability, if by parametrization we mean a relation between the properties of the cumulus ensemble and the large-scale variables at the same instant." (Arakawa and Schubert 1974, 691). The point is that the collective behavior of the cumulus ensemble can be considered to vary (quasi) instantaneously with the large scale system, and as a consequence, the small scale processes that occur in every single cumulus cloud are negligible.

We have seen how the scale separation assumption (both in space, where the cumulus ensemble occupies a small area compared to the whole system ($\sigma_c \ll 1$), and in time, where $\tau_{ADJ} \ll \tau_{LS}$) is at the basis of the physical argument of quasi equilibrium of the cloud ensemble with the large scale process that allows scientists to ignore small scale cloud processes in a dynamical description of the large scale process. In so far as the cloud ensemble maintains properties that are in quasi equilibrium with the large scale, smaller scale cloud physics will not have any direct

[39] These parameters, too, represent smaller scale cloud processes and their interaction with the environment. For example, the vertical mass flux in a cloud $\widehat{M}$ depends on smaller scale processes within the cloud. These parametrizations also rely assumptions about what processes become relevant at what spatiotemporal scales.

relevance for changes in the large scale. As a consequence, the scale separation assumption contributes to the theory to phenomena inference in so far as it isolates the large scale phenomenon of interest from irrelevant physics at smaller scales.

4.0 STRUCTURAL UNCERTAINTY: TARGET SYSTEM IDENTIFICATION

4.1 INTRODUCTION

So far, I have illustrated the importance of the scale existence assumption and the scale separation assumption for the identification of phenomena as targets. The discussion so far has touched on issues of idealization and abstraction. I will now show that my approach of analyzing philosophical issues from the perspective of target system identification can be helpful for discussing another philosophical issue: the nature of uncertainty in scientific modeling.

Uncertainty is an aspect of climate science that has drawn increasingly more attention in recent years (see, for example, Parker 2006, 2010, 2011; Frigg et al. 2013, 2014; Stainforth, Allen, et al. 2007; Stainforth, Downing, et al. 2007; Knutti 2008). Uncertainty also is an aspect of climate science that is particularly important epistemically and socio-politically. On one hand, there is a limited and problematic sense in which experiments can be performed to validate and confirm our models.[40] On the other hand, the climate directly affects individuals and society world-wide, and taking the changes in climates and our current knowledge concerning the climate into account is an important part of policy-making, both at national and international levels.

[40] See the debate between Parker (2009a) and Lloyd (2010) and the papers by Steele and Werndl (2013, 2015) and references therein. For general issues concerning simulations as experiments see Parker (2009b) and Morrison (2009).

The institution that takes care of the socio-political and scientific interface is the Intergovernmental Panel on Climate Change (IPCC). One of their tasks is to summarize the state of the science and communicate it to policymakers in the "Summary for Policymakers." In these summaries, uncertainty is usually communicated in terms of likelihoods of events occurring and confidence in statements about the climate. For example, the guidance note for lead authors of the IPCC Fifth Assessment Report introduces confidence in the following way:

> Confidence in the validity of a finding is based on the type, amount, quality, and consistency of evidence (e.g., data, mechanistic understanding, theory, models, expert judgment) and the degree of agreement. (Mastrandrea et al. 2010, 1)

This statement identifies many different sources of uncertainty, mostly related to data and to the models used to obtain prediction. The sources of uncertainty have been discussed by philosophers and scientists alike, and they be broadly categorized as data uncertainty, parameter uncertainty and structural uncertainty (Parker 2010). Data uncertainty is uncertainty about the reliability of data collection and analysis methods, and uncertainty about whether this data is accurately representing the relevant state of the climate when it is fed into large numerical models (also called General Circulation Models). Parameter uncertainty is uncertainty about the reliability of the empirical values assigned to the parameters in a model. Finally, structural uncertainty is uncertainty about the structure of the equations of a model. The structure of the model is understood, in this case, as the number of mathematical objects that represent the target and the relation between these objects. This chapter will focus on structural uncertainty.

There are several important reasons to be interested in uncertainty in general, and in structural uncertainty in particular. Scientists and philosophers describe structural uncertainty as one of the types of uncertainty that is particularly hard to assess (see Knutti 2008; Stainforth, Allen, et al. 2007, Stainforth, Downing, et al. 2007; Parker 2010, 2011; Frigg et al. 2013, 2014). So, one is interested in asking whether there is a way of characterizing structural uncertainty in order to obtain reliable uncertainty assessments. Further, structural uncertainty is an epistemic, representational problem. In particular, philosophers are interested in asking how the construction of a model is epistemically justified and the extent to which models can accurately represent their targets. Last, policy makers, scientists and philosophers are interested in articulating an account of structural uncertainty that allow for better communication across the different disciplines. This point has been clearly made by Smith and Stern (2011) and by Stainforth, Downing, et al.:

> A two-way communication between climate scientists and users of climate science is . . . of fundamental importance. Only by understanding the needs of different [policy] sectors can the science be usefully directed and communicated. Only by understanding the *conditions, assumptions and uncertainties* of model based statements about future climate can decision makers evaluate the relevance of the information and make informed, if subjective, assessments of risk. (Stainforth, Downing, et al. 2007, 2165, my emphasis.)

Stainforth, Downing, et al. claim that policy makers and scientists need to be able to clearly state and assess the uncertainties that arise in models. These uncertainties are tied to the conditions under which model based statements are made.[41] One such condition is the range for which a model can be used to obtain reliable predictions—be these various emission scenarios, parameters values, etc. Uncertainties are also tied to the assumptions of the model based statements, which are the idealizing and abstracting assumptions that scientists introduce in their models. An account

[41] Model based statements can be either predictive statements about the future of the earth's climate or statements that reflect the scientists' understanding of the physics of the climate.

of structural uncertainty that highlights the sources of uncertainty would allow for confident decision making even without entirely eliminating the sources of uncertainty. Part of the challenge of formulating such an account is to individuate the epistemic and representational constraints that are tied to the conditions and assumptions of model based statements.

In this chapter, I will start by reviewing some of the existing literature on structural uncertainty. I will argue that while philosophers have identified an important problem, their approach misses some crucial points. Next, I will develop a set of desiderata aimed at addressing the gaps in the current literature. These desiderata are also aimed at guiding the development of an account of structural uncertainty that is useful to scientists, philosophers and policymakers. Finally, I will start sketching an account of structural uncertainty that addresses these desiderata. This account will be approached by starting from the assumptions that are involved in identifying targets and building models of the targets.

4.2 CURRENT ACCOUNTS

Two philosophers that have been discussing structural uncertainty explicitly are Frigg et al. (2014) and Parker (2006, 2010, 2011). Their accounts are representative of most of the scientific literature on the topic. Reviewing these accounts will serve the purpose of clarifying what we mean by structural uncertainty and highlight the deficiencies of the current debate on the topic.

Frigg and his colleagues develop the discussion of structural uncertainty in the context of nonlinear models. In order to focus on structural uncertainty, they assume that the modeler has perfect knowledge of parameter values and no uncertainty concerning the collection, analysis and input of data into computer models. In their view:

> A model has [structural uncertainty] if the model dynamics differ from the dynamics in the target system. If a nonlinear model has only the slightest [structural uncertainty], then its ability to generate decision-relevant probability predictions is compromised. (Frigg et al. 2014, 32)

This statement implies that for non-linear systems, even slightest change in model structure can compromise ability to generate decision-relevant predictions. The authors call this the "hawkmoth effect" (Frigg et al. 2014, 39). This effect is similar to sensitive dependence to initial conditions for chaotic systems: as Frigg et al. say, for both cases, it does not matter how close the model structure or the initial data is to the structure or data that will allow the model to make accurate predictions—small deviations in model structure or data will lead to predictions that are misleading.

A question that can be raised of this account is whether there are cases for which this effect can be managed, or whether there is a way of determining whether we are in a regime where this effect can be severe or not. Frigg et al. are pessimistic about being able to be able to obtain a non-arbitrary measure on a class of models that would allow scientists to determine which models are trustworthy and which are not–their suggestion is to discard probabilities as meaningful tools for assessing predictions (Frigg et al. 2014, 57). However, they overlook the target identification and model building process in their discussion.

Another account that is more directly addressing practical issues in climate modeling is provided by Parker (2006, 2010, 2011). Parker says, "Structural uncertainty often refers to uncertainty about the form that modeling equations should take" (Parker 2010, 265). This means that structural uncertainty is uncertainty about the number of mathematical objects that need to be taken into account in order to accurately model a target. To clarify her account of structural uncertainty, Parker points out several sources of structural uncertainty. In particular, structural uncertainty arises when

> i. Physical processes of interest are not described by well-established theories.
>
> ii. The representation of physical processes involves simplifications and idealizations, because of theoretical or pragmatic constraints.
>
> iii. Processes need to be parametrized, and there is no best way to parametrize a process. (Parker 2010, 264–265)

As in the account provided by Frigg et al., structural uncertainty in Parker's case is an unavoidable aspect of modeling. Simplifications, idealizations and parametrizations are part and parcel of mathematically modeling the climate. A question that arises for the use of simplifications and idealization is whether they are introduced in such a way that uncertainty can be mitigated. What is the role of idealizations and simplifications for obtaining the model-based statements mentioned by Stainforth et al.'s quote in the previous section? In order to understand how uncertainty is tied to these modeling practices, one needs a more in-depth discussion of what limitations these idealizations and simplifications introduce.

Further, the Navier-Stokes equations (which contain some parameters themselves) are not analytically solvable.[42] This implies that the spatiotemporal grid on which these equations are solved numerically also requires that further parametrizations be introduced in the model. In order to understand the importance of parametrizations in models, a more detailed description of the role of parametrization in models is needed.[43] A parameter, in general, can be understood as a black

[42] The Navier-Stokes equations are a set of fluid dynamical equations that lie at the basis of modeling of the atmosphere. Climate models involve a lot more than the use of these equations.
[43] A full account of the use of parametrizations in mathematical modeling would be an extremely worthwhile project but is beyond the scope of this paper. For a brief overview, see Edwards (2010, Ch.13).

box that incorporates finer detail of the processes represented in the model. This finer detail is usually thought to not matter much, and to the extent it does, it is represented as an effective parameter the value of which can usually be measured empirically. For example, consider the Navier-Stokes momentum equation:

$$\rho \frac{Du}{Dt} = -\nabla \bar{p} + \mu \nabla^2 u + \frac{1}{3} \mu \nabla (\nabla \cdot \mathrm{u}) + \rho g$$

In this equation, the parameters are ρ, p and μ. These are density, pressure and viscosity respectively. Each parameter represents a type of molecular motion and/or interaction that can, for the purpose of modeling fluid, be represented by the cumulative behavior of the molecules[44]. In other words, certain fine details of molecular motion are not relevant for the description of their cumulative behavior, and the parameters suppress these details into one term. In the case of the Navier-Stokes equations, these parameters are extremely effective.[45] It would be therefore too quick to assert that parameters introduce uncertainty. Rather, the challenge that arises to assess when parameters introduce uncertainty in models is to understand when certain details can be considered irrelevant, and when instead they need to be taken into account.

In order to address some of the lacunae that exist in Frigg et al.'s and Parker's accounts, the assumption that we can take the model as a given needs to be left behind. The way simplifications, idealizations and parametrizations contribute to uncertainty depends on how they are introduced in models. They also are an indication of the domain of applicability of the model. The applicability of a model depends, as a matter of fact, on its intended target, and as a consequence, focusing on how targets are identified is a promising strategy to investigate how

[44] An interesting related debate concerns reductionism. See Kim (1992, 1999) and Batterman (2009) for an overview of the arguments for and against the possibility of reductionism in science.
[45] The Navier-Stokes equations are currently the best tool for modeling fluids.

uncertainty can enter the modeling process. As I have argued in chapters 2 and 3, important assumptions that are involved in the target identification process are about the scale at which the target occurs, and whether targets occurring at different scales are spatiotemporally separated enough to be considered phenomena that can be modeled independently of one another. The use of these assumptions is crucial in order to identify whether details at other scales can indeed be parametrized or not. To give an account of structural uncertainty that thoroughly reflects the modeling practice, thus, we need to focus on how targets are identified and the scale dependence of targets.

4.3 DESIDERATA

An adequate account of structural uncertainty should not only address the scientific motivations, such as being able to manage and assess uncertainty but should also address the philosophical motivations (How is any particular model epistemically justified?) and socio-political motivations (How can scientists and policy makers communicate effectively about uncertainty?) that drive the interest in providing such an account. One of the reasons that drives this interest is to develop an account of structural uncertainty that is useful—an account that will allow the interested party to provide a clear assessment of this kind of uncertainty. In this section I will describe three desiderata that reflect these motivations.

Epistemic Reliability. An adequate account of structural uncertainty should indicate when a model is epistemically reliable. This desideratum addresses both the scientific and the philosophical motivations. There are two main ways in which epistemic reliability can be specified. These are

i. Specification of the conditions under which the model makes accurate predictions.

ii. Indication of when the model is accurately representing target phenomena.

The first item concerns the extent to which a model is applicable to the world. For example, a model that is used for daily weather prediction is not going to be reliable—will be very uncertain––for making monthly forecasts, and vice versa. The second item concerns representational accuracy: an adequate account of structural uncertainty should indicate when phenomena are properly identified and how thoroughly and perspicuously the model manifests the phenomena and their interactions. In other words, representational accuracy captures the extent to which a model represents the regularities that we observe in the world and whether the modeler has an adequate explanation for it.

Sources of Uncertainty. An adequate account of structural uncertainty should reveal sources of uncertainty in the model. This criterion addresses the philosophical motivations for such an account: it should highlight which aspects of the model construction and justification process are likely to introduce uncertainty. Two ways in which sources of structural uncertainty can be revealed are:

i. Shedding light on the assumptions made in constructing the model that generate structural uncertainty.

ii. Indicating the domains of inapplicability given the assumptions.

The first item concerns the modeling construction and justification process: a good account of uncertainty should identify what can introduce uncertainty. The use of modeling assumptions is directly tied to Parker's point about the use of simplifications and idealizations. Some simplifying and idealizing assumptions are going to be conducive to building a reliable model, while some will not. The second item is related to how the use of various assumptions can constrain the applicability of a model. For example, given some simplifying and/or idealizing assumption, a model will be reliable for one domain or purpose but not reliable for another. In other words, this is a criterion that requires an adequate account of structural uncertainty to provide an explanation for the unreliability of the model.

Policy Relevance. Last, an adequate account of structural uncertainty should indicate how structural uncertainty can vary for different questions asked by the policy makers. As Stainforth et al. suggest, the identification of the sources of uncertainty is a two-way process, from the scientist to the policy maker and from the policy maker to the scientist. A clear communication of the expectations from and the limits of models should make the uncertainty identification process clearer. Expectations from models may concern the kind of phenomena, spatial and/or temporal scales that the model should predict reliably. Limits of models similarly concern the limits on the range of predictions that can reliably be made with those models. Identifying these expectations and limits does not mean that uncertainty will be eliminated. Nevertheless, an account of structural uncertainty that meets these criteria can reduce the so-called "unknown unknowns" and shed more light on the "known unknowns". Reducing the number of unknown unknowns should help the policy maker differentiate between the cases for which a sound assessment of risk can be carried out and when there just is not enough information to make such an assessment.

The criteria that I have identified are the guide for developing an adequate account of uncertainty. I understand that these represent more an ideal towards which an account should strive rather than criteria that are easily attainable. Nevertheless, these criteria are best addressed when studying scientific (mathematical) modeling from the perspective of target identification. Understanding the assumptions that are involved in selecting targets and building models thereof will help the philosopher develop an account of when a model can be trusted (*epistemic reliability*), why a model cannot be trusted (*sources of uncertainty*) and how structural uncertainty depends on questions asked by the policy makers (*policy relevance*).

4.4 A NEW ACCOUNT

I will now show that by taking into account how assumptions about scales are used in identifying target systems and constructing models, we can start providing an account of structural uncertainty that meets the needs of scientists, philosophers and policy makers.

In climate science, we want to identify phenomena, model their behavior, and model the way they interact. As I have already discussed in this dissertation, phenomena relevant for climate modeling occur at many different scales. This point has been made clearly by the meteorologist Kerry Emanuel (1986).[46] While this is an important aspect of climate science, there are good reasons to believe that this is a problem that is not confined to this discipline. For example, I have shown in chapter 1 that Levin (1992) has emphasized the importance of the concept of scale for identifying ecological phenomena. Loeb and Imara (2017) have recently made a similar case for astrophysics.

4.4.1 Parametrizations

An example of the scale dependence of phenomena in climate science, as I have discussed in chapter 3, arises for the case of modeling hurricanes. Scientists are interested in modeling their formation, development, trajectories and their dissipation into tropical storms. At the same time, we are interested in modeling smaller scale phenomena, like clouds–which can be modeled

[46] See chapter 2.

individually or as parts of hurricanes. Hurricanes and clouds are described by different physics, yet they are both important for climate modeling. This is part of the more general issue of identifying and representing phenomena at many different scales, and incorporating the different physics dominates at those different scales. In climate science phenomena occur across 10 order of magnitude, from 10s of meters to the entire globe, and from seconds to centuries.

The task of identifying phenomena at different scales, modeling their physics and the interaction of these phenomena at different scales is particularly complex. One of the most important goals of climate scientists is to make predictions about how climate phenomena will change over time and how changes in phenomena at different spatiotemporal scales affect one another: for example, scientists are interested in knowing how global average temperature increase (a large-scale phenomenon) will affect smaller scale phenomena like hurricane frequency and intensity in the Caribbean (Emanuel 1999, 2005). For this reason, understanding scale dependence and scale interaction is crucial for obtaining reliable predictions.

As mentioned above, climate models are numerical models that are solved by large computers. In order to solve numerical models, the equations of the models that describe how phenomena of interest will change over time will need to be discretized. To complicate matters further, the way the discretization is carried out also depends on the phenomena we are interested in. For example, if we want to predict how global sea level increase will affect tidal increase in the long term, we will chose a scale of discretization that is small enough to capture the important information about tidal increase but ignores oceanic phenomena at smaller spatiotemporal scales like tsunamis.[47] In general, when placing out models on a spatiotemporal grid so that we can use

[47] Tsunamis only temporarily increase sea levels and related tides, whereas the sea level increase due to the melting of ice has much more long term effects on the height of tides.

computers to make predictions, the scale of the grid is a function of the characteristic scale of the phenomenon of interest. [48]

A question that follows naturally from this is: why can we ignore details at smaller scales when we are interested in large scale phenomena? And why don't we always use the most fine-grained model? A first answer has already been hinted at in chapter three. In discussing the role of parametrization schemes, I mentioned the remarks on this topic by the atmospheric scientist Arakawa. For easy reference, here is the important passage quoted before:

> Even under a hypothetical situation in which we have a model that resolves all scales, it alone does not automatically give us an understanding of scale interactions. Understanding inevitably requires simplifications, including various levels of "parametrizations" . . . Parametrizations thus have their own scientific merits. (Arakawa 2004, 2496)

Here, Arakawa stresses the importance of understanding how phenomena occurring at different scales behave and how they interact with phenomena at other scales. In order to do this, we need to isolate phenomena from their environment, which is accomplished by ignoring those parts of the world that are considered to be, to a certain extent, irrelevant for the description of the phenomenon—a task performed by parameters. As a consequence, one of the reasons we do not use the most fine-grained model is that we need to isolate phenomena of interest, and to do this we use parametrizations. [49] An important corollary of this statement is that, contra Parker,

[48] Note that the computational grid scale is not always one-to-one with the REV that is used to obtain parametrizations. The relation between parametrization and computational grid is an interesting issue that contributes to structural uncertainty in significant ways. The epistemic significance of the relation between parametrization and computational grid is discussed, among other topics, in Wilson (2017). I thank Mark Wilson for drawing my attention to this issue.

[49] I have argued for this claim extensively in chapters 2 and 3.

parametrizations are not always problematic: in many cases they can reduce structural uncertainty by providing theoretical insight.

In fact, when parametrizations are used to isolate phenomena, they highlight the relevant components of a target system by suppressing the irrelevant details occurring at other scales. By highlighting the relevant components of a phenomenon, parametrizations are a tool for advancing theoretical understanding of the phenomena, and reduce uncertainty that is tied to lack of such understanding.

To fully appreciate Arakawa's point, it is helpful to further look into the process of parametrization in climate science. One of the most important underlying assumptions of the parametrization process is the scale separation assumption. This is the assumption that says that if two processes occur at spatial and temporal scales that are sufficiently separated, then they can be modeled separately. This is the assumption that underlies the identification of target systems at different scales: finding situations where the scale separation assumption holds is a central part of identifying phenomena and constructing tractable models (Baldissera Pacchetti 2017). I would like to note that the scale separation assumption is neither a sufficient or necessary condition for target system identification, but an important widespread heuristic.

Understanding the central role of the scale separation assumption in climate modeling offers a more satisfactory account of structural uncertainty: when the scale separation assumption is not clearly applicable, the exclusion of details via parametrization schemes is not well justified. The lack of a satisfactory justification for a parametrization scheme leads to structural uncertainty as we do not know if the mathematical object used to model an aspect of the world does so accurately. In what follows, I am going to discuss the philosophical implications for structural uncertainty of the Arakawa and Schubert parametrization scheme I have discussed in chapter 3.

For the sake of clarity, I will briefly repeat the most salient details of the physics of that case. I will then illustrate how understanding the role of this assumption offers a more satisfactory account of structural uncertainty.

When scientists want to isolate phenomena like hurricanes, they want to know whether they need to model the microphysics of individual clouds within a hurricane or whether they can ignore the individual clouds and model the cloud behavior collectively. Further, scientists are interested in how the collection of clouds interacts with its macroenvironment–for example, a hurricane. Groups of clouds are pretty stable structures in hurricanes, so scientists can model the clouds as a group, understand this stable collective behavior, and, while we are at it, save some computing power.

One important way to model the stability of a group of clouds within a larger structure is to relate the statistical effects of the cloud microphysical processes to the macrophysical ones: the processes that occur within a cloud to the collective changes that occur in the group of clouds and the changes in the environment. The microphysical processes will be negligible if the changes in the physics of individual clouds occur at scales that are much smaller than the scale of the system they are part of. This assumption is the scale separation assumption: if there is enough scale separation between microphysical and macrophysical processes, then—if we are interested in modeling macrophysical processes—the microphysics can be ignored: the microphysics of the cloud is considered to be irrelevant for the physics of the interaction of the cloud with the macroenvironment.

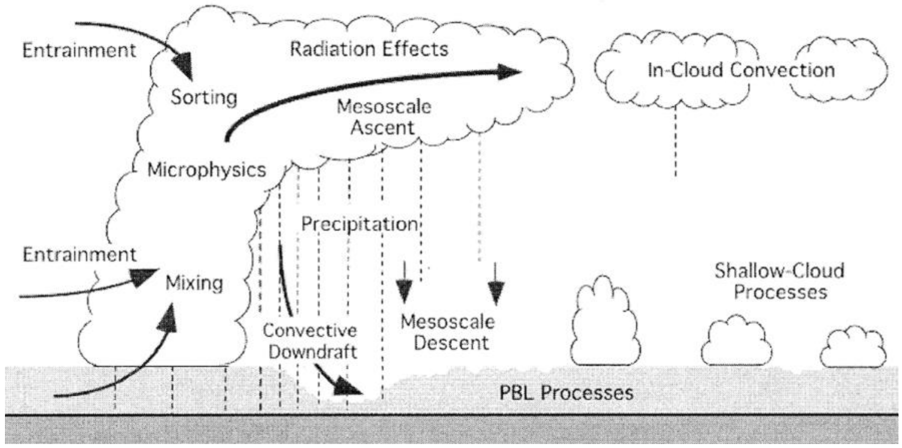

Figure 4.1. Cloud and associated processes for which major uncertainties in formulation exist. From Arakawa (2004). ©American Meteorological Society. Used with permission.

The scale separation assumption is not just a mathematical construct that allows the scientist to eliminate some degrees of freedom from their models, but it also involves a physical justification of the physical processes involved. For the case discussed here, the physics that gets ignored are the processes that occur internally to each cloud, such as the conversion of humidity into rain droplets. What instead is related to the macroenvironment is the entrainment—collection of energy from the environment in the form of hot air and moisture—and the release of energy in the form of precipitation, as represented in Figure 4.1. These are the processes that are in balance with the macrophysics. The physical interpretation of the scale separation is that the internal microphysics can be considered to adjust virtually instantaneously to the changes in the macroenvironment. The temporal scale separation between the macrophysical and microphysical processes is so great, that in effect we can treat these two processes as being in balance. At this particular scale, when we "see" groups of clouds behaving this way, we can model the larger scale system as a closed dynamical system and effectively ignore the physics that occurs at smaller scales, because this physics does not matter for the description of this large-scale system.

One should note, however, that this parametrization process is not only guided by theoretical and empirical considerations: Sometimes, the parametrization is guided by the size of the grid necessary to make our models numerically solvable. What guides parametrization in general is the relation between processes at different scales and their interaction. A smaller scale process is parameterized when it is not explicitly resolved in our climate models.

4.4.2 Structural Uncertainty

We have seen so far that the scale separation assumption is one of the key ingredients in justifying parametrization schemes used to isolate phenomena. As a consequence, whether a parametrization scheme introduces structural uncertainty in the model will partly depend on whether the use the scale separation assumption is warranted. If the justification for ignoring the physics is not valid, then we do not know whether we have made a legitimate move in model building. It also means that we do not know whether our system will need the explicit representation of the physics at smaller scales. From this starting point, we can begin providing an account of structural uncertainty that will meet the desiderata spelled out in section 4.3.

Structural uncertainty is uncertainty about whether the structure of the mathematical model accurately represents its target. Given role of the scale separation assumption in the Arakawa and Schubert parametrization scheme, we can say that structural uncertainty arises when one or more of the following conditions obtain:

i. The scale separation assumption cannot be justified theoretically.
ii. The scale separation assumption cannot be verified empirically.
iii. Complex models do not reliably account for inter-scale interactions between different "closed system" models of phenomena.

Let me elaborate on these conditions. As I argued above, the application of the scale separation assumption is not just a mathematical construct used to eliminate degrees of freedom. In many cases, the use of this assumption is accompanied by a physical interpretation of the processes that are modeled explicitly, and the ones that are suppressed in a parameter. In the case of the parametrization described above, one can eliminate the microphysics of the cloud because, at the scale at which the cloud ensemble is modeled, the internal processes of the clouds are in balance with their surroundings. Further, the use of this assumption should be backed by data series: the separation in scale must be observable from an analysis of data collected at the scales of interest.[50]

The assumption also plays an important role for the mathematical structure of the model: it provides a "closure condition" for the set of equations that describe the target system.[51] This role is what effectively separates the target system at the scale at which it is modeled from the details from the details at other scales. As Arakawa notes, having a closed set of equations will help scientists gain theoretical insight into the behavior of an ensemble of clouds. If the set of equations does so successfully, then it serves as a stepping stone for understanding the interaction of the ensemble of clouds with smaller and larger scale phenomena. On the other hand, lack of such understanding of inter-scale interactions can lead to structural uncertainty–as the interactions between different phenomena cannot be modeled reliably (or do not have a proper justification).

---

[50] Yano and Plant (2012) have argued that scale separation for the time series of cloud precipitation cannot be verified empirically and problematize the concept of quasi-equilibrium as a physical justification for the Arakawa-Schubert parametrization scheme described above. This is a case where uncertainty can be highlighted by problematizing the use of the scale separation assumption.

[51] Closure conditions, in mathematics, are conditions that allow for the solution of a set of equations so that the number of free variables does not exceed the number of equations that are solved simultaneously.

This issue is particularly important in climate science, as phenomena can be observed to occur on many scales.

4.4.3 Desiderata

So far, I have argued that by paying attention to the identification of target systems and the assumptions that are involved in modeling the target we can start providing a more detailed account of when and where structural uncertainty can arise in the modeling process. For this account to be useful, however, it needs to satisfy the desiderata I have developed in section 4.3.

First of all, an adequate account of structural uncertainty should indicate when a model is epistemically reliable. In order to accomplish this, it needs to specify the conditions under which the model makes accurate predictions and accurately represents the target phenomena. My account satisfies this criterion by looking at the conditions under which structural uncertainty arises with respect to how the scale separation assumption is used to identify and model phenomena at particular scales. For example, when we have empirical validation and a theoretical interpretation of the use of the scale separation assumption, we can say that a model of the target is epistemically reliable.[52] The use of the scale separation assumption can also serve as a platform for starting to answer more detailed questions about the reliability of a model. Scale separation has consequences for the conditions under which a model can make accurate predictions, for example: the model will only make accurate predictions for the scales at which phenomena are represented. One condition for which a model is accurately representing the target system is if the relevant scales are taken into account.

[52] This is equivalent to a non vicious convergence of the theory to phenomena inference and data to phenomena inference as described in chapter 2.

Second, an adequate account of structural uncertainty should reveal the sources of uncertainty in the model. Studying structural uncertainty from the perspective of target identification and model building allows the philosopher and the scientist to shed light on the assumptions made in constructing the model that generate structural uncertainty. The scale separation assumption is one such assumption: it plays a crucial role in eliminating irrelevant details from the model. As a consequence, an incorrect application of this assumption is a source of uncertainty. The content of this assumption also indicates the domains of inapplicability of a particular model given the assumptions. For example, the parameters that are introduced in a model will be a constraint on the kind of processes that are predicted by the model and the predictions that cannot be made by the model. Since the assumption effectively ignores details at smaller scales, reliably downscaling a particular model to be able to reliably represent smaller scale processes will not be an easy task. In order to downscale reliably, scientists will need to identify what phenomena are relevant at various smaller scales and how they interact with phenomena at other scales. A downscaling process that does not take theoretical constraints of the scale assumption into consideration will end up being unreliable (see, for example, the analysis by Frigg et al. (2013) of a worrying case of downscaling).

Third, an adequate account of structural uncertainty should address the needs of policy makers by indicating how structural uncertainty can vary for different policy related questions. For example, if we only need infrastructure that will withstand average rainfall over a large spatial region, the model on which we base our decisions can reliably ignore the short-distance causes of the rain accumulation, e.g. whether it is due to constant steady rainfall or extreme weather events (e.g. hurricanes). If, on the other hand, we need to protect a coastal area from extreme climate events, that model has considerable structural uncertainty about these phenomena and is not useful for making policy decision. In sum, because of the details of the role of scale assumptions in modeling, this is an account of structural uncertainty that is useful in a way that existing accounts are not.

4.5 CONCLUSION

In this chapter, I have argued that philosophers have identified an important problem in climate science; namely, how best to assess and communicate the nature of structural uncertainty in climate modeling. However, their characterization of structural uncertainty is not sufficiently well-explicated. I have argued that a useful account of structural uncertainty should meet desiderata that are driven by scientific, philosophical and policy motivations. By focusing on the context of identifying phenomena and model building and highlighting the centrality of assumptions about scale separation, I have provided an account of structural uncertainty that addresses epistemic and representational problems facing scientists and philosophers as well as the socio-political needs of policy makers. This is not a complete account, but it highlights how a different strategy for answering philosophical questions can provide insights into scientific practice that have not been highlighted enough so far.

5.0 LEVELS AND SPATIOTEMPORAL SCALES

5.1 INTRODUCTION

In the previous chapters, I have shown how target system identification, especially for the case of climate science, relies on assumptions about the scale at which phenomena occur. We have seen that these assumptions are empirical assumptions, and they are used to isolate relevant aspects of the world from irrelevant ones in order to build models. In other words, scale related assumptions are used by scientists as a tool to partition the world into phenomena that can inform scientists about the structure of the world. One way in which philosophers have also talked about the structure of the world and the way scientists describe and organize it is in terms of so-called "levels." Levels are introduced to describe how scientist construct successful models about the world in spite of its complexity (Wimsatt 2007; Mitchell 2003, 2009) and how scientist use abstraction to achieve different modeling goals (Levins 1966, 1993, 2006). On these views, levels capture particular aspects of the world that constitute target systems for models, in such a way to be conducive to good modeling practices, whatever the measure for "goodness." Levins (1966, 1993, 2006) and Wimsatt (2007) mainly engage in a descriptive analysis of the concept of levels in the sciences. Sometimes this concept is also part of normative accounts (e.g., Mitchell 2003).[53] In this

[53] The debate about levels extends far beyond these authors, especially in the philosophy of biology. See for example the debate on reductionism in biology (Bringandt and Love 2017) or on

chapter, I focus on the role of the concept of level for models as described by these philosophers. The discussion of levels in this framework has both ontological and epistemic implications.

5.2 CURRENT ACCOUNTS

5.2.1 Levins on Levels

Levins equates the notion of level to that of abstraction. Abstraction, in this case, is a process that is supposed to capture those aspects of the world that make up the target system chosen by the researcher.

The clearest formulation of this concept can be found in Levins' response to a critique of Orzak and Sober (1993). In the passage below, Levins is discussing the relation between his criterion of realism, the system of interest, the level of abstraction and the assumptions in a model:

> [T]he more closely the *assumptions of a model* correspond to the *processes and level of abstraction* being studied, the more realistic a model is; and the more closely the characteristics of interest correspond to the outcome of the model, the more realistic the model is. (Levins 1993, 549; my emphasis)

This passage can be read in the following way. At different "levels of abstraction," one can identify different processes, and for each of these processes a model can be more or less realistic depending on the assumptions made by the model. In this reading, a level of abstraction identifies a process to be studied, i.e. the target system that is modeled.

---

the units and levels of selection in evolutionary biology (Lloyd 2017). I will, however, rely on the authors cited in the main text as these are the most general contributions, and hence most pertinent to the purpose of the present essay.

Levins' characterization of the relation between the assumptions of a model and the level of abstraction resonates with the role of the scale related assumptions in identifying targets that has been discussed in the previous chapters. As I have argued in chapter 2, the scale related assumptions are *abstracting* assumptions. Assumptions about the scale of a target system guide scientists in determining the regularities that make up a phenomenon. In other words, these assumptions isolate the system of interest from irrelevant details–they are part of the assumptions that allow scientists to identify the level of abstraction that corresponds to a particular phenomenon. The level of abstraction described by Levins and the spatiotemporal scale at which a system is identified described in my work so far are therefore similar concepts.

Moreover, Levins claims that there is no straightforward, monotonic relation between number of independent variables in a mathematical model and the realism of the model. As a matter of fact, realism can be *reduced* "when the added variables or connections among variables change the level of abstraction of a model" (Levins 1993, 549). Therefore, it is not always correct to say that a model that has more variables is more realistic than a model with fewer variables. The addition and subtraction of variables and the relations among them can have different effects in mathematical models, and this is dependent on the level of abstraction the model is supposed to capture. Variables can sometimes be added to include more representational detail. Relations between variables can be modified to represent processes differently. [54] These two modifications to mathematical models can increase their realism. These modifications can, however, in some cases, decrease realism. The cases in which this happens is when the modification of the model

[54] An example of modification of relation between variables is changing a linear equation to a non-linear one.

deviates from the model’s intended level of abstraction, i.e. when the model does not represent the process captured by that level of abstraction.

Again, this resonates with the role of scale related assumptions in justifying particular parametrization schemes used in modeling phenomena. As I have argued in chapters 3 and 4, suppressing details is not always a pragmatic move that subsequently needs to be corrected by explicitly modeling those details. If a particular parametrization scheme provides the closure for a model in such a way that the relevant components of a phenomenon are isolated, then, in Levins’ terms, the model is realistically modeling the phenomenon. Adding more variables from processes at, say, smaller scales would change the phenomenon and the level for which one is constructing a model.

The context of application of the model also plays a role in determining the extent to which a model represents its target system. As a consequence, any ambiguity in interpreting Levins' concept of “level of abstraction” comes from at least two sources. The first source is his rejection of the analysis of realism of a model in terms of necessary and sufficient conditions. The second source is the tension between the representational aspect of the model and the context in which the model is used. Levins’ concept of abstraction is in fact constrained both by the modeler’s interest (the context) and by the world (representational aspect). Levins’ murkiness on the issue of model representation is probably intentional. He claims that “formal analysis prefers to work with properties that belong to the object in itself, independent of its context” (Levins 1993, 549).[55] He claims that there are different ways in which model representation can add to considerations about

[55] By mentioning formal analysis, Levins is referring to the standard analytical method employed by philosophers of finding necessary and sufficient conditions that their objects of analysis must satisfy. In this case, Levins is arguing against such an analysis of the realism desideratum.

mathematical models, and that formal analysis has limited the extent to which such considerations can be fruitful for philosophical analysis.

As I read Levins, the ambiguity with respect to abstraction can be resolved in the following way: abstraction can be seen as capturing two different aspects of modeling. One aspect identifies the process in the world to be studied. In this sense, we abstract away from the rest of the world, and focus only on the system of interest. The other identifies the assumptions that will allow the scientist to decide which variables and parameters describing the system of interest will serve the purpose she is building the model for. This reading clarifies our understanding of the realism desideratum: the scientist does not strive towards a unique, all comprehensive model when she strives for realism. Rather, a model becomes more and more realistic as it captures the relevant features of the world that make the process of interest a target system as such.

Another insight into Levins' analysis of abstraction is found in his distinction between abstractions of perspective and level. Levins' levels and "abstraction of perspective" (Levins 2006, 744) are in line with my interpretation given above. The "perspective" (Levins 2006, 744) is the initial question asked by the scientist (the context). The level of abstraction is the aspect of the world that has to be considered given one perspective (the representational aspect of the model in its context). Depending on the level, the system exhibits characteristic properties. These properties are captured in the model by means of dynamic variables and constant parameters (Levins 2006, 744), which together with the context in which the model is applied to, contribute to the realism of the model.

Table 5.1. Abstractions of perspective, extent and level in fruit-fly ecology. Reprinted with permission from Springer Customer Service Centre GmbH: Springer Nature, Biology and Philosophy, Strategies of Abstraction, R. Levin, ©2006.

Perspective	Horizontal scale	Temporal scale	Dynamics	Constants
Temperature tolerance	Individual fly	Minutes to hours	Mortality	Fly biology, temperature
Adaptation to temperature	Individual fly	Days to a week	Growth and development, acclimation	Fly biology, temperature regimes
Behavior in relation to temperature	Population of flies of one species	Minutes	Attraction to food versus heat stress	Habitat pattern of temperature, food resources
Demography	Population of flies of one species	Seasonal	Reproduction versus mortality	Habitat, community of species
Community	Ecosystem of interacting species	Months to years	Competition, predation	Habitat, community of species
Micro-evolutionary	Single species	Years	Natural selection versus migration and drift	Habitat, community of species

Table 5.1 summarizes the abstractions of perspective and level introduced by Levins. The leftmost column describes the perspectives: these are the general questions that scientists can ask about objects in the world. The remaining columns capture the level: given a particular question, the levels constrain what can be considered as a target system, what is directly relevant to it, and what is irrelevant. For example, if a scientist is interested in the temperature tolerance of a species, e.g. the fly, the relevant horizontal (spatial) scale is that individual fly, and its mortality rate in a constant temperature environment over minutes to hours (the temporal scale). The details of the biology of the fly are not directly relevant and hence are represented by a constant parameter.

Levels of abstraction pick out the relevant target systems in the world that occur at characteristic spatiotemporal scales.

5.2.2 Wimsatt on Levels

Wimsatt also makes an explicit connection between levels and target systems, reifying levels and emphasizing their role in epistemic issues in mathematical modeling. Levels capture the "structural features . . . that dominate our world" (Wimsatt 2007, 194). The concept of level is fundamental in describing the complex ontology of the world. Entities that are found at a specific level, and the levels themselves, are seen as "real objects" (Wimsatt 2007, 195). Levels are characterized in terms of the entities that are found at each level, in terms of the relations (usually causal) between the entities, and their composition. Wimsatt says:

> Levels…are constituted by families of entities usually of comparable size and dynamical properties, which characteristically interact primarily with one another, and which, taken together, give an apparent rough closure over a range of phenomena and regularities. (Wimsatt 2007, 203–204)

The size of the entities found at each level is an important factor in the characterization of the level themselves, as he agrees with Haldane's (1926) claim that size plays an essential role in determining which physical processes are at play between the entities at a given level (Wimsatt 2007, 207). For Wimsatt, size refers to spatial extension of the entities and object at each level, and their surface to volume ratio (Wimsatt 2007, 207). The size of the entities is related to their dynamical properties.

The importance of the dynamical processes in characterizing a level of organization implies that time also plays a very important role in determining levels (Wimsatt 2007, 216). "Closure," in the quotation above, means that a theory (or class of models) aimed at describing one level can successfully describe all the entities that arise at that level. In this context, entities at a level and their interactions can be interpreted as target systems: entities or phenomena the properties of which can be measured and recognized at a particular level constitute the target system for which models are built in order to describe the dynamics of such entities or phenomena.

5.2.3 Mitchell on Levels

Mitchell (2003, 2009) talks about levels indirectly in her discussions of the metaphysical considerations in and epistemological issues of building models of a complex world. Her main goal is to provide a model of scientific enterprise that involves a "critical pluralism" of models and perspectives (Mitchell 2003, 4). Critical pluralism can be understood on one hand as an ontological pluralism: for one target system different models of aspects of properties of the target system can be formulated, and they all capture relevant processes of the target system. None of these models is the preferred one.

Mitchell is criticizing an assumption Kim's (1999) account of reductionism and emergence is committed to, namely that there is a preferred, fundamental level of description. She says:

> What the philosophical arguments that assume unique, complete representations and a privileged level and straightforward mappings from one level to another miss entirely is how is the property at the higher level is produced, and what are the difference among the many kinds of relationships between higher- and lower-level properties that occur in nature? (Mitchell 2009, 32)

Levels, on Mitchell's account, are stable, material levels of organization, where higher levels are composed of lower levels, and (mostly) present a higher degree of complexity depending to the nature of the interaction of the lower level parts (such as positive or negative feedbacks, nonlinear interactions, etc.). Levels form a "hierarchy of multiple levels of organizations in biological systems" (Mitchell 2009, 21). An individual flying starling, for example, is a lower level component of a flock of starlings that moves in organized and complex patterns.

Mitchell's discussion is framed in terms of the debate about reduction and emergence. While I do not intend to engage with this debate, there are several important points about models, representation, and their relation to levels that emerge from Mitchell's analysis. Requiring unique, complete representations is not reasonable, as representations are partial by nature (Mitchell 2009, 31). The partiality of representations is tied to the existence of levels of organizations: complex phenomena have properties that do not exist at the lower level, and such properties cannot be predicted by lower levels.

While Mitchell does not explicitly connect levels and phenomena (as a matter of fact, she seems to imply that a phenomenon can be described at many levels of complexity), it nevertheless seems reasonable to suggest that, in her framework, target systems are constrained by levels: stable properties or structures at different levels indicate what entities can be associated with particular properties or processes.[56]

In this section I have described how Levins, Wimsatt, and Mitchell use the concept of level (of abstraction and of organization) to analyze the part of the world that is represented in models. All three philosophers agree that these levels of organization are an aspect of the world that allows

[56] See Mitchell's discussion of natural selection acting at different levels of complexity (2009, 36)

scientists to identify entities and their properties in order to construct models. While Levins and Wimsatt do suggest that spatiotemporal scales play a role in defining scale, they do not develop the role in much depth. Further, Mitchell only refers to a hierarchical organization of levels, in which lower level entities make up the higher level entities. I suggest that in order to rigorously analyze the way scientists identify target systems, the notion of spatiotemporal scale has to be seen as *central* to an account that aims to clarify the constraints that are imposed by the world on models (or on scientists constructing models). The clearest connection between the concept of level and the role of scale related assumptions can be made for Levins' account. In the next section, I will argue for such a connection in general terms.

5.3 LEVELS AS SPATIOTEMPORAL SCALES

In the previous section, I have sketched how some philosophers have argued that nature presents scientists with various levels of organization, at which and across which various description of nature can be provided through models. The views on levels provided by these philosophers share some characteristics. These characteristics are that levels exist in the world and as a consequence they impose constraints on how target systems are identified for successful modeling. Further, these levels are organized in terms of complexity for Mitchell, and (at least partially) in terms of scales for Levins and Wimsatt. I suggest that focusing on the role that spatiotemporal scales play for identifying target systems can provide a more accurate analysis of what philosophers have called "levels" and the role they play in model building.

As has emerged from the examples discussed in this dissertation, processes are described in terms of stable patterns that are observed in the world, and patterns are easily identifiable when they occur at characteristic spatiotemporal scales. Scientists associate target systems with processes that occur at these scales, such as tsunamis, hurricanes, or the daily migration of zoo-plankton from depths to the sea surface (see Figures 1.1 and 1.2). *The identification of target systems involves identifying the characteristic scales at which patterns can be observed,* and the scale existence and scale separation assumptions are important tools used in this process of identification. In addition, scientist have to define properties that are associated with a measurable quantity (e.g. motion and kinetic energy or population density and biomass variability), and they have to detect the signal of the pattern despite "the contamination of the other parts of the spectrum" (Stommel 1963, 573).

This is where the notion of level can be clarified by the analysis of the role of the scale related assumptions that I have provided in the previous chapters: levels can be clearly individuated when there is enough scale separation that the pattern can be distinguished from the noise coming from other scales. The constraining role of the world that philosophers have attributed to levels comes from spatiotemporal scales. On the other hand, when many spatiotemporal scales interact, target systems are harder to identify, and pragmatic components dominate the target identification process. In this perspective on modeling, the patterns occurring at characteristic spatiotemporal scales are the aspect of the world constraining the creative aspect of modeling in science. Scale separation of various processes can be seen in Figure 1.1 and Figure 1.2. Peaks of the relevant quantities of interest (biomass variability, kinetic energy) at various spatiotemporal scales correspond to processes that can be isolated from their environment. The assumption that systems can be isolated—that one can appeal to scale separation to identify the system—is an idealization. There will always be *some* interaction across these scales, even if modeling a system involves setting up boundaries for the system.

This characterization of levels also gives force to Mitchell's analysis of levels discussed in section 5.2.3: philosophical arguments that assume a unique, complete representation of the world and that there is a privileged level of description are misguided. They miss how properties at different levels are related to one another, and this is a very important aspect of scientific practice: when scientists isolate phenomena at their characteristic scales, they only focus on the scale of interest. The privileged level is only the level at which the phenomenon of interest is identified and modeled.

As mentioned above, the process of identifying characteristic scales can be tricky where phenomena that occur on different scales interact and when phenomena have components occurring at many scales. See, for example, the spatiotemporal scales of tidal terms and meteorological effects in Figure 1.2. These two processes occur at overlapping temporal and spatial scales, which means some systems in this range will show properties that are a combination of both tidal terms and meteorological effects[57]. The mesoscale, described by Emanuel (1986)—see section 1.3—, also presents the scientists with similar problems: a smooth energy spectrum might mean either that phenomena occurring at different but contiguous scales interact, and that phenomena could be identified but it is hard to do so, or it might mean that there are no phenomena of interest at all. In this case, pragmatic considerations are at the forefront of the identification of phenomena at particular scales and the modeling process. From a theoretical perspective, the validity of the scale separation assumption means that these processes can be described in terms of stabilities of the system, i.e. in terms of those factors that restore the system's balance of forces.

[57] Floods can be the effect of the interactions of phenomena at different scales. A case from meteorology that occurs rather frequently is the combination of the action of winds (meteorological effects) and supermoon tide (when tidal terms are stronger due to the closeness of the moon to the earth) on the height of the ocean surface.

Once the system can be identified and isolated from its environment, modeling the system involves identifying direct variables, time-varying parameters and constant parameters. These elements describe and quantify the patterns - turning them from observations into models of target systems.

The use of the scale related assumptions indicates an ontological commitment to target system as processes confined to characteristic spatiotemporal scales. This ontological commitment drives the epistemic consequences of the use of the scale existence and scale separation assumptions. As I have argued in chapter 4, the use of these assumptions in constructing models constrains the predictive ability of the model and the structural uncertainty associated with such predictions. Since patterns arise at characteristic spatial and temporal scales, the predictive ability of the model will only work on the sales at which these patterns arise (i.e. the model for this pattern will not be able to provide predictions on longer or shorter temporal scales, or finer or coarser spatial scales, mainly because these scales are outside of the scope of the model)[58]. Further, if scientists want to generate predictions for phenomena on longer time scales and/or finer spatial scales (these are the challenges that climate scientists are facing today with general circulation models), then there needs to be an integration of systems and their interaction *across* the scales of interest.

These points imply that identifying systems at their characteristic spatiotemporal scales is central for both qualitative understanding of and quantitative predictions from models[59]. Understanding how target system are isolated from their environment and the way models capture

---

[58] This statement agrees with Mitchell's claim that a characteristic of emergence is that the emergent patterns are not predictable from its components at lower levels taken individually (2009, 35). It must be noted that this unpredictability also goes the other way around: the behavior of the parts, if taken in isolation, cannot be predicted from the behavior of the whole.

[59] A similar point has been made by Potochnik and McGill (2012). While I do not agree with their criticism of the literature on levels, I do agree with the general positive argument about spatiotemporal scales.

variables of interest at their characteristic scales is therefore at the center of the modeling practice. These points further support the arguments presented in chapter 2 that scale assumptions are idealizing while also being empirical.

5.4 CONCLUSION

In this chapter I have argued that by taking the perspective of the identification of target systems, "levels" can be best understood in terms of spatiotemporal scales at which characteristic patterns are observed and associated with phenomena. I have analyzed the views of Levins, Wimsatt and Mitchell, and showed that their characterization of levels is compatible with the account I have given of the role of the scale assumptions. These assumptions are abstracting tools that scientists use to individuate phenomena, and as such they characterize levels in Levins' sense. Wimsatt, on the other hand, characterizes levels as capturing dominant structural features in the world. He explicitly makes the connection between these dominant structures of the world and the "size" of these structures, making the connection between the role of the scale assumptions and levels particularly clear. Mitchell, on the other hand, argues for an important feature of levels that is also characteristic of the assumptions about scales: there is no privileged level and the "mappings" between levels are not at all straightforward. This is in line with my argument that details that are not relevant for modeling phenomena at their characteristic scales are ignored *during the modeling process* and given the complex nature of this process and of the world itself, adding those details back in is not easy at all. I have especially argued for this point in chapter 4: climate scientists face considerable difficulties and can introduce structural uncertainty in their models when they downscale their models, which is equivalent to “adding more details” about phenomena at lower levels. Focusing on assumptions about scales can therefore provide a characterization of levels that is true to modeling practice and connects some of the most influential arguments in the literature on levels.

6.0 CONCLUSION

Modelers don't aim at obtaining a total description of the world, but rather partial representations of the world through models. The general philosophical question that is asked in this context is how these partial representations can produce reliable knowledge about the world. In particular, philosophers have asked how reliable knowledge can be obtained despite the fact that abstractions and idealizations, usually thought of as distortions of the world, are introduced in scientific models.

In this dissertation I have argued that an important that scientists use in their representations are assumptions about the scale at which phenomena occur and the scale related conditions that allow scientists to treat phenomena as separate from one another. Assumptions about spatiotemporal scales play an important role in determining what counts as a target system: in chapter 2 I argued that these assumptions are ineliminable abstracting and idealizing assumptions. These assumptions are ineliminable as they play an essential role in determining what counts as a phenomenon, and removing these assumptions would imply that the scientists are not representing the intended phenomenon in their models. Removing such assumptions, would imply that scientists changing the level at which phenomena are being modeled, given the relation between scale assumptions and levels illustrated in chapter 5. The empirical—and sometimes normative—justification for the use of these assumptions allows scientists to assess the reliability of the representation of a phenomenon, as I argued in chapter 4. When such justifications are absent, we have what is called in the literature "structural uncertainty". This is uncertainty about whether the structure of a mathematical equation accurately reflects the intended phenomenon. Chapter 3 provided an extended discussion of the role of scale in the individuation of phenomena both from theoretical principles—the Arakawa and Schubert parametrization scheme discussed in section 3.3—, and from data—the identification of ENSO discussed in section 3.2.

My approach has been to focus on an aspect of scientific practice that is at the heart of scientific modeling, namely the identification of target systems. This approach distinguishes itself from traditional philosophical discussions of models, where typical questions focus on how models can provide reliable knowledge of the world despite their being idealizing and abstracting. I have focused on how relevant variables and parameters for the target system are chosen. This approach has allowed me to argue that the scale assumptions used in choosing relevant variables and parameters are necessarily idealizing and abstracting. Because these assumptions are ineliminable, it is *in virtue of* these idealizations and abstractions that scientists can obtain reliable partial representations of the worlds—the models of phenomena.

I have also contributed to the conversation about the distinction between data, phenomena and theory that had been sketched by Bogen and Woodward (1988) and Woodward (2011). In particular, I have taken their distinction to be the first stepping stone needed in discussing the identification of targets and the role of scale related assumptions therein. In this dissertation, I showed that the relation between spatiotemporal scales and relevant properties of the system is a first important step in analyzing how phenomena are isolated from the rest of the world, both from observations and from theory.

The empirical nature of the assumptions can further inform us about the nature of the phenomena that these assumptions help identify. Spatiotemporal scales are constraints on the kind of regularities that can be identified. As Levin says (Levin 1992, 1947), scientists can move from observing irregular patterns at one scale to regular patterns at other scales, and these regular patterns allow for the kind of generalizations that allow for insight into the behavior of phenomena in the world. Phenomena are therefore not just the product of the modeler's purpose or some arbitrary normative judgment. Phenomena are stable patterns in the world. The patterns in the world are identifiable at different spatiotemporal scales. By taking the perspective of target system identification, "levels" can be best understood in terms of spatiotemporal scales at which characteristic patterns are observed and associated with phenomena. In other words, phenomena present themselves at different scales, rather than different "levels" as often described in the literature.

Depending on the scale a scientist is interested in, different kinds of detail will need to be included in the model. For example, modeling a mesoscale hurricane will not require that the small scale details of the molecules that the hurricane is made out of be included in the model—and in some cases not even the details of the single clouds need to be included in such a model: the scale separation assumption in this case guides scientists in determining when such details need to be included or can be excluded.

This role of assumptions about scales in identifying phenomena is also an important indicator of when structural uncertainty—uncertainty about whether the model accurately represents a phenomenon—can arise in climate science. When scientists are not justified in leaving out or including certain details from the world in their descriptions of phenomena, then this uncertainty arises.

These conclusions show that studying the process of target system identification is an approach that is not without interest for philosophers of science. This study in fact not only highlights what kind of assumptions are introduced when identifying phenomena and building models, but it also provides insight into the nature of models and their relation to the world. Models are in fact constrained by normative assumptions and the purposes of modelers, but they are also importantly constrained by the patterns that arise at different scales. This constraint is embodied by the use of the scale existence and scale separation assumptions in scientific modeling. The process of identifying target systems is therefore not a process that is beyond the scope of philosophical analysis, but rather a process that can provide important insight both in classic and emerging discussions in the philosophy of science.

BIBLIOGRAPHY

Arakawa, A. 2004. “The Cumulus Parameterization Problem: Past, Present, and Future.” *Journal of Climate* 17 (13): 2493–2525.

———, and W. H. Schubert. 1974. “Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part I.” *Journal of the Atmospheric Sciences* 31 (3): 674–701.

Baldissera Pacchetti, M. 2017. “A Role for Spatiotemporal Scales in Modeling.” *Studies in History and Philosophy of Science* 67: 14–21.

Barnston, A. G., M. Chelliah, and S. B. Goldenberg. 1997. “Documentation of a Highly ENSO-Related SST Region in the Equatorial Pacific.” *Atmosphere-Ocean* 35 (3): 367–383.

Batterman, R. W. 2001. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press.

———. 2002. “Asymptotics and the Role of Minimal Models.” *The British Journal for the Philosophy of Science* 53 (1): 21–38.

———. 2009. “Idealization and Modeling.” *Synthese* 169 (3): 427–446.

———. 2013. “The Tyranny of Scales.” In *The Oxford Handbook of Philosophy of Physics*, edited by R. W. Batterman, 255–286. New York: Oxford University Press.

Bjerknes, J. 1969. “Atmospheric Teleconnections from the Equatorial Pacific.” *Monthly Weather Review* 97 (3): 163–172.

Bogen, J., and J. F. Woodward. 1988. “Saving the Phenomena.” *The Philosophical Review* 97 (3): 303–352.

———. 1992. “Observations, Theories and the Evolution of the Human Spirit.” *Philosophy of Science* 59 (4): 590–611.

———. 2005. “Evading the IRS.” In *Idealization XII: Correcting the Model. Idealization and Abstraction in the Sciences*, edited by M. R. Jones and N. Cartwright, 233–268. Amsterdam: Editions Rodopi.

Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Davis, C. A., and K. A. Emanuel. 1991. “Potential Vorticity Diagnostics of Cyclogenesis.” *Monthly Weather Review* 119 (8): 1929–1953.

Edwards, P. N. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: MIT Press.

Elliott-Graves, A. 2014. “The Role of Target Systems in Scientific Practice.” PhD diss., University of Pennsylvania. ProQuest (UMI No. 3635498).

Emanuel, K. A. 1986. “Overview and Definition of Mesoscale Meteorology.” In *Mesoscale Meteorology and Forecasting,* edited by P. S. Ray, 1–17. Boston, MA: American Meteorological Society.

———. 1999. “Thermodynamic Control of Hurricane Intensity.” *Nature* 401: 665–669.

———. 2005. “Increasing Destructiveness of Tropical Cyclones Over the Past 30 Years.” *Nature* 436: 686–688.

———. 2007. “Quasi-Equilibrium Dynamics of the Tropical Atmosphere.” In *The Global Circulation of the Atmosphere*, edited by T. Schneider and A. H. Sobel, 186–218. Princeton, NJ: Princeton University Press.

Frigg, R., L. A. Smith, and D. A. Stainforth. 2013. “The Myopia of Imperfect Climate Models: The Case of UKCP09.” *Philosophy of Science* 80 (5): 886–897.

Frigg, R., S. Bradley, H. Du, and L. A. Smith. 2014. “Laplace’s Demon and the Adventures of His Apprentices.” *Philosophy of Science* 81 (1): 31–59.

Giere, R. 1988. “Laws, Theories, and Generalizations.” In *The Limitations of Deductivism*, edited by A. Grünbaum and W. C. Salmon, 37–46. Berkeley, CA: University of California Press.

Gray, A. R., and S. C. Riser. 2015. “A Method for Multiscale Optimal Analysis with Application to Argo Data.” *Journal of Geophysical Research Oceans* 120, 4340–4356.

Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Oldenborgh, and T. Stockdale. 2009. “Understanding El Niño in Ocean–Atmosphere General Circulation Models: Progress and Challenges.” *Bulletin of the American Meteorological Society* 90: 325–340.

Haldane, J. B. S. 1926. “On Being the Right Size.” *Harper’s Magazine* 152: 424–427.

Haury, L. R., McGowan, J. A., & Wiebe, P. H. 1978. “Patterns and Processes in the Time-Space Scales of Plankton Distributions.” In *Spatial pattern in plankton communities*, edited by J. H. Steele, 277–327. Springer, Boston, MA.

Holling, C. S. 1992. “Cross-Scale Morphology, Geometry, and Dynamics of Ecosystems.” *Ecological Monographs* 62 (4): 447–502.

Houze Jr., R. A. 1993. *Cloud Dynamics*. San Diego, CA: Academic Press.

Jochum, M., and R. Murtugudde, eds. 2006. *Physical Oceanography: Developments Since 1950*. New York: Springer.

Jones, M. R. 2005. “Idealization and Abstraction: A framework.” In *Idealization XII: Correcting the Model. Idealization and Abstraction in the Sciences,* edited by M. R. Jones and N Cartwright, 173–217. Amsterdam: Editions Rodopi.

Katz, R. W. 2002. “Sir Gilbert Walker and a Connection Between El Niño and Statistics.” *Statistical Science* 17 (1): 97–112.

Kim, J. 1992. “Multiple Realization and the Metaphysics of Reduction.” *Philosophy and Phenomenological Research* 52 (1): 1–26.

———. 1999. “Making Sense of Emergence.” *Philosophical Studies* 95: 3–36.

Knutti, R. 2008. “Should We Believe Model Predictions of Future Climate Change?” *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 366 (1885): 4647–4664.

Laymon, R. 1985. “Idealizations and the Testing of Theories by Experimentation.” In *Observation, Experiment, and Hypothesis in Modern Physical Science,* edited by P. Achinstein and O. Hannaway, 147–173. Cambridge, MA: MIT Press.

Levin, S. A. 1992. “The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture.” *Ecology* 73 (6): 1943–1967.

Levins, R. 1966. “The Strategy of Model Building in Population Biology.” *American Scientist* 54 (4): 421–431.

———. 1993. “A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science.” *The Quarterly Review of Biology* 68 (4): 547–555.

———. 2006. “Strategies of Abstraction.” *Biology and Philosophy* 21 (5): 741–755.

Lloyd, E. A. 2010. “Confirmation and Robustness of Climate Models.” *Philosophy of Science 77*: 971–984.

———. 2017. “Units and Levels of Selection.” In *The Stanford Encyclopedia of Philosophy (Summer 2017 Edition)*, edited by E. N. Zalta. URL = <https://plato.stanford.edu/archives/sum2017/entries/selection-units/>.

Loeb, A., and N. Imara. 2017. “Astrophysical Russian Dolls.” *Nature Astronomy* 1: 0006.

Brigandt, I., and A. Love. 2017. “Reductionism in Biology.” In *The Stanford Encyclopedia of Philosophy (Spring 2017 Edition)*, edited by E. N. Zalta. URL = <https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/>

Mastrandrea, M. D., C. B. Field, T. F. Stocker, O. Edenhofer, K. L. Ebi, D. J. Frame, H. Held, E. Kriegler, K. J. Mach, P. R. Matschoss, G.-K. Plattner, G. W. Yohe, and F. W. Zwiers. 2010. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. Intergovernmental Panel on Climate Change (IPCC). URL = <http://www.ipcc.ch>.

McAllister, J. W. 1997. “Phenomena and Patterns in Data Sets.” *Erkenntnis* 47 (2): 217–228.

McMullin, E. 1985. “Galilean Idealization.” *Studies in History and Philosophy of Science* 16 (3): 247–273.

McPhaden, M. J. 2006. “El Niño and Ocean Observations: A Personal History.” In *Physical Oceanography: Development Since 1950*, edited by M. Jochum and R. Murtugudde, 79–99. New York: Springer.

McPhaden, M. J., A. J. Busalacchi, R. Cheney, J. R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith, and

K. Takeuchi. 1998. “The Tropical Ocean-Global Atmosphere Observing System: A Decade of Progress.” *Journal of Geophysical Research* 103 (C7): 14169–14240.

Mitchell, S. D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

———. 2009. *Unsimple Truths: Science, Complexity, and Policy*. Chicago: University of Chicago Press.

Montgomery, M. T., and R. K. Smith. 2014. “Paradigms for Tropical Cyclone Intensification.” *Australian Meteorological and Oceanographic Journal* 64 (1): 37–66.

Morgan, M. S., and M. Morrison, eds. 1999. *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Morrison, M. 2009. “Models, Measurement and Computer Simulation: The Changing Face of Experimentation.” *Philosophical Studies* 143 (1): 33–57.

Orzack, S. H., and E. Sober. 1993. “A Critical Assessment of Levins’s ‘The Strategy of Model Building in Population Biology’ (1966).” *The Quarterly Review of Biology* 68 (4): 533–546.

Parker, W. S. 2006. “Understanding Pluralism in Climate Modeling.” *Foundations of Science* 11 (4): 349–368.

———. 2009a. “Confirmation and Adequacy-for-Purpose in Climate Modelling.” *Aristotelian Society Supplementary Volume* 83: 233–249.

———. 2009b. “Does Matter Really Matter? Computer Simulations, Experiments, and Materiality.” *Synthese* 169 (3): 483–496.

———. 2010. “Predicting Weather and Climate: Uncertainty, Ensembles and Probability.” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 41 (3): 263–272.

———. 2011. “When Climate Models Agree: The Significance of Robust Model Predictions.” *Philosophy of Science* 78 (4): 579–600.

Pedlosky, J. 1987. *Geophysical Fluid Dynamics*. New York: Springer-Verlag.

Peschard, I. 2012a. “Modeling and Experimenting.” In *Models, Simulations, and Representations,* edited by P. Humphreys and C. Imbert, 42–61. New York: Routledge.

———. 2012b. “Forging Model/World Relations: Relevance and Reliability.” *Philosophy of Science* 79 (5): 749–760.

Pincock, C. 2009. “From Sunspots to the Southern Oscillation: Confirming Models of Large-Scale Phenomena in Meteorology.” *Studies in History and Philosophy of Science* 40 (1): 45–56.

———. 2012. *Mathematics and Scientific Representation*. Oxford: Oxford University Press.

Potochnik, A., & McGill, B. 2012. The limitations of hierarchical organization. *Philosophy of Science 79* (1): 120-140.

Smith, L. A., and N. Stern. 2011. “Uncertainty in Science and Its Role in Climate Policy.” *Philosophical Transactions of the Royal Society A* 369: 4818–4841.

Steele, K., and C. Werndl. 2013. “Climate Models, Calibration, and Confirmation.” *The British Journal for the Philosophy of Science* 64: 609–635.

———. 2018. “Model-Selection Theory: The Need for a More Nuanced Picture of Use-Novelty and Double-Counting.” *The British Journal for the Philosophy of Science* 69: 351–375.

Stainforth, D. A., M. R. Allen, E. R. Tredger, and L. A. Smith. 2007. "Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions." *Philosophical Transactions of the Royal Society of London A* 365: 2145–2161.

Stainforth, D. A., Downing, T. E., Washington, R., Lopez, A., & New, M. 2007. "Issues in the Interpretation of Climate Model Ensembles to Inform Decisions." *Philosophical Transactions of the Royal Society of London A* 365: 2163–2177.

Stommel, H. 1963. "Varieties of Oceanographic Experience." *Science* 139: 572–576.

Strevens, M. 2004. "The Causal and Unification Approaches to Explanation Unified—Causally." *Noûs* 38: 154–176.

Trenberth, K. E. 1997. "The Definition of El Niño." *Bulletin of the American Meteorological Society* 78 (12): 2771–2777.

Weisberg, M. 2013. *Simulation and Similarity: Using Models to Understand the World*. New York: Oxford University Press.

Wimsatt, W. C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Wilson, M. 2017. *Physics Avoidance: and other essays in conceptual strategy*. Oxford University Press.

Woodward, J. F. 1989. "Data and Phenomena." *Synthese* 79 (3): 393–472.

———. 2000. "Data, Phenomena, and Reliability." *Philosophy of Science* 67: S163–S179.

———. 2011. "Data and Phenomena: A Restatement and Defense." *Synthese* 182: 165–179.

Yano, J.-I., and R. S. Plant. 2012. "Convective Quasi-Equilibrium." *Reviews of Geophysics* 50: RG4004.