THE UNIVERSITY OF CHICAGO

STATISTICAL METHODS FOR CLIMATIC PROCESSES WITH TEMPORAL NON-STATIONARITY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF STATISTICS

BY
ANDREW NAMAN POPPICK

CHICAGO, ILLINOIS
JUNE 2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank my advisor, Michael Stein, for his support for my intellectual growth since my first undergraduate course in statistics and throughout my graduate education, and for the inspiring example he sets for rigorous and self-scrutinized applied statistical research. This work would likewise not have been possible without the support and guidance of Elisabeth Moyer, who adeptly upholds the ideals of relevant and well-communicated interdisciplinary science.

I additionally thank my committee members, Mihai Anitescu and Matthew Stephens, for their attention to this dissertation and role in my education. Through the department's consulting program, Peter McCullagh and Mei Wang have also provided me invaluable opportunities to be engaged in applied statistics and have been a continued source of useful feedback. Finally, I have benefitted in many ways from conversations with and support from a number of my peers and colleagues; I would particularly like to thank Maryclare Griffin, Guillaume Pouliot, Michael Horrell, Scott Powers, Matthew Reimherr, Eric Janofsky, Somak Dutta, Dinah Shender, and Walter Dempsey.

For the work comprising Chapter 2, I thank William Leeds and Joseph Guinness for helpful discussions and for providing relevant code, and David McInerney for providing the climate model runs that I analyze. Support for this work was provided by STATMOS, the Research Network for Statistical Methods for Atmospheric and Oceanic Sciences (NSF-DMS awards 1106862, 1106974, and 1107046), and RDCEP, the University of Chicago Center for Robust Decision Making in Climate and Energy Policy (NSF grant SES-0951576). This work was completed in part with resources provided by the University of Chicago Research Computing Center. For the work comprising Chapter 3, Rao Kotamarthi and Scott Collis at Argonne National Laboratory provided a helpful discussion; this work was partially supported by STATMOS, and by the U.S. Department of Energy (grant DE-SC0002557). Work in this dissertation has appeared in slightly different form in *The Annals of Applied Statistics* (Poppick et al., 2016) and *Environmetrics* (Poppick and Stein, 2014).

ABSTRACT

Three chapters are included in this dissertation. Chapter 1 introduces the research herein.

The societal impacts of future climate change depend on changes in temperature variability in addition to changes in mean temperatures. While general circulation models (GCMs) predict changes in both means and variability, GCMs cannot fully reproduce present-day temperature distributions. In Chapter 2, we address an ensuing need for simulations of future temperatures that combine the observational record and GCM projections of changes in means and variability. Our perspective is that such simulations should be based on transforming observations to account for GCM projected changes, in contrast transforming GCM output to correct for discrepancies with observations. Our methodology is designed for simulating transient (nonstationary) climates, which are evolving in response to changes in CO_2 concentrations (as is the Earth at present). Since the proposed simulation requires GCM projected changes in covariance, we describe a statistical model for the evolution of temporal covariances in a GCM, and apply this model to an ensemble of runs from one GCM, CCSM3. We find that, in CCSM3, changes in covariance can be explained as a function of the regional mean change in temperature and the rate of change of warming. This feature means that our statistical model can be used to emulate the evolving covariances in the GCM under scenarios for which the GCM has not been run. When combined with an emulator for mean temperature, our methodology can simulate temperatures under such scenarios. The emulator for variability changes is also of interest on its own as a summary of GCM projections of variability changes.

Chapter 3 concerns modeling a bivariate meteorological process – temperature and dew point – measured at high temporal frequencies and given covariate information. Dew point is bounded above by temperature, so nonstandard methods are needed to characterize their bivariate distribution, especially at times when temperature approaches dew point (or equivalently when relative humidity approaches 100%). The data analyzed are minute-to-minute measurements from early May from the years 2003 through 2012 at the Atmospheric Radi-

ation Measurement (ARM) Program's Southern Great Plains (SGP) site in Northern Oklahoma, at the central facility near Lamont, OK. We find that, in addition to relative humidity, solar radiation and the magnitude of minute-to-minute changes in wind direction help explain aspects of the behavior of the bivariate process. We propose a parametric model for how the spectral matrix of a high-frequency component of the process varies with these covariates over time. The spectral approach allows for convenient and interpretable models of bivariate processes in time and our model captures many of observed features of the data.

CHAPTER 1

INTRODUCTION

There are two major components of this dissertation. These components are unified by the exploitation of time-varying spectral methods for modeling and simulating temporally nonstationary atmospheric processes and the use of covariates to model and understand temporal nonstationarity.

1.1 Overview of Chapter 2

In Chapter 2, I propose a method for simulating future temperatures whose means and temporal covariances are evolving in time in response to greenhouse gas forcing. Simulations of future temperatures may be required, for example, as inputs into studies that assess the societal impacts of climate change. The societal impacts of climate change depend on changes in temperature variability – not just mean temperature – but changes in variability are not yet well understood. The primary tools used for studying changes in the distribution of climate variables are atmosphere-ocean general circulations models (GCMs). These are complex physical models that can generate runs of modeled climate under specified forcing scenarios. GCM runs provide useful information about climate change, but they should not be used on their own as inputs in impacts studies because they do not reproduce observed climate distributions under historical forcing scenarios, a fact well-understood in the climate community. There is therefore a need for methods that combine GCMs and observations to produce future simulations that preserve features of the observed climate and capture changes in means and variability projected by GCMs.

Popular approaches for generating such simulations generally fall into two categories. A *model-driven method* applies a transformation to GCM output that attempts to correct features of the GCM runs compared to observations; an *observation-driven method* instead transforms observations in a way that captures GCM projected changes (e.g., Ho et al.

(2012), although the above terminology is our own). An advantage of observation-driven methods is that, unlike model-driven methods, they preserve features of the observations not explicitly accounted for in the procedure. The full spatial-temporal distribution of observed temperatures is complicated, and projected changes in the distribution of temperatures are not very large compared to discrepancies with GCMs, so, in our perspective, observation-driven methods are more likely than model-driven methods to produce realistic simulations. Most existing observation-driven methods, however, only account for changes in means and marginal variances projected by GCMs. Changes in variability importantly depend on timescale (e.g., day-to-day fluctuations vs. those at seasonal or inter-annual scales), so changes in marginal variances do not completely summarize changes in variability.

We propose an observation-driven simulation method for temperatures in future transient climates that accounts for GCM projected changes in both means and temporal covariances. The proposed simulation method exploits the representation of locally stationary processes in terms of time-varying spectra (Priestley, 1981; Dahlhaus, 2012) to alter the observations temporal covariance structure to reflect changes projected by the GCM. The crucial fact that makes the procedure useable is that it is possible to transform an observed time series into one with a new mean and covariance structure using only information about how its mean and covariance structure change in the future. This is important because GCMs are assumed more informative about changes than absolute levels. The proposed simulation requires no modeling of the observations and preserves many features of the original observations; of the GCM, my method requires only a statistical model for its projected changes in mean and covariance structure.

Modeling GCM projected changes in temporal covariances in a transient climate is itself a challenging statistical task. We find that, in the GCM runs studied, one can characterize projected changes in covariances as a function of mean warming and the rate of change of warming. In so doing, our statistical model not only can be used for the proposed simulation procedure, but it also provides an interpretable summary of the GCM variability changes, of

interest on its own to the climate community. We show that the statistical model performs well as an emulator for the GCM projected variability changes: it may be used to predict the GCM projected variability changes in future forcing scenarios under which the GCM has not been run. This is important because, since they are so computationally expensive, GCMs are usually run under fewer future scenarios than are of interest for impacts assessments. In the GCM studied, temperature variability is projected to decrease as temperatures increase, particularly at shorter timescales and in the mid-latitudes, and a stronger change in variability (holding mean warming fixed) is projected in climates where warming is occurring more quickly.

1.2 Overview of Chapter 3

In Chapter 3, I discuss the problem of modeling a bivariate meteorological process – temperature and dew point – measured at high temporal frequencies and given covariate information. Since dew point is a lower bound for temperature, standard methods based on Gaussian processes cannot be used directly to model such a bivariate process. The high frequency nature of the data presents its own curiosities: the frequent measurements allow one to see clearly that the process is nonstationary over the course of the day and across days. In the data studied, both variables exhibit more variability, for example, during daylight hours, when relative humidity is lower (i.e., when temperature is not close to dew point), and when the wind direction is changing rapidly.

In the proposed model, the bivariate temporal spectral matrix of a residual process evolves as a function of covariates and is constrained so that the original process essentially respects the physical bound that temperature be above dew point. Modeling the process in the spectral domain allows one to easily write bivariate models that are both valid and interpretable, and capture many of the subtle features apparent in the data.

CHAPTER 2

TEMPERATURES IN TRANSIENT CLIMATES: IMPROVED METHODS FOR SIMULATIONS WITH EVOLVING TEMPORAL COVARIANCES

2.1 Introduction

Assessing the potential impacts of future climate change on areas of societal interest, such as agriculture and public heath, requires an understanding of how climate features important to those areas are expected to change. Impacts often depend on more than just the response of global or even local mean temperatures to greenhouse gas forcing. Many agricultural crops, for example, are highly sensitive to even brief periods of stress temperatures, particularly at certain times of the growing cycle, so crop yields can be strongly affected by changes in temperature variability even in the absence of a change in mean (e.g., Wheeler et al. (2000)). In part because of examples like this, the climate and impacts communities have been interested in understanding changes in temperature variability in future climates.

Potential future changes in temperature variability are not yet well understood. By its third assessment report, the Intergovernmental Panel on Climate Change (IPCC) stated that there was some empirical evidence for a decrease in variability at intra-annual timescales, but sparse evidence for changes in inter-annual variability (IPCC, 2001). More recent studies have not produced more definitive conclusions, with results apparently depending on specific definitions of variability and timescale as well as on the region being studied (IPCC, 2007). (The most recent report, IPCC (2013), frames variability changes in the context of extreme events, which are not a subject of this work.) One example of a physical mechanism that might explain variability changes at intra-annual timescales in a particular region is that changes in the polar jet stream can produce more persistent weather patterns over, for example, North America (e.g., Francis and Vavrus (2012)), but the mechanism and even

detection of these changes remains controversial (Screen and Simmonds, 2013; Barnes, 2013). Implicit in the broader discussion about variability is that variability changes can differ by timescale of variation or geographic location. Because specific impacts will depend on timescale and geographic location, methods for assessing changes should be able to make such distinctions. That is, understanding projected changes in variability relevant to impacts is a problem of understanding the changes in covariance structure of a spatial-temporal field that is evolving in time.

Beyond empirical studies, the primary tools used to understand and project changes in the distribution of climate variables are atmosphere-ocean general circulation models (GCMs). GCMs are deterministic, physical models that are used to generate runs of modeled climate under, for example, varying forcing scenarios. While GCMs are deterministic, the climate system being modeled is chaotic and so GCM realizations under the same forcing scenario but with different initial conditions will behave as if they were statistically independent. Summarizing the statistical properties of GCM predictions under different forcing scenarios is a challenge on its own.

That said, GCM runs alone are often not sufficient as inputs for impacts assessments, which may require realistic simulations from the full distribution of the Earth's temperatures. It is well understood that GCMs somewhat misrepresent observed temperature distributions under present conditions: regional mean temperatures may differ from observations by several degrees, and studies have noted discrepancies between higher order moments of the modeled and observed climate distributions (IPCC (2013) and references therein). On the other hand, GCMs are assumed to produce informative projections of, for example, future changes in mean temperatures due to changes in greenhouse gas forcing: the underlying physics are relatively realistic, and GCMs are able to reproduce observed temperature trends in historical forcing runs (e.g., IPCC (2013)). Projections of variability changes also have many consistent features across different GCMs, although current studies do not address changes in full covariance structures (e.g., Schneider et al. (2015); Holmes et al. (2015)).

Impacts assessments researchers have therefore recognized a need to understand not only how temperature distributions are changing in GCM runs, but also how to combine those projections with the observational record to produce simulations of temperatures that more likely follow the distribution of real future temperatures.

There are two popular classes of approaches for generating such simulations: those that modify GCM output to account for model-observation discrepancies (model-driven procedures), and those that modify observational data to account for changes projected by GCMs (observation-driven procedures); see, for example, Ho et al. (2012) and Hawkins et al. (2013) for reviews of common strategies. The most basic model-driven procedure is simple "bias correction," where the difference in mean between observed temperatures and those in historical GCM runs is estimated and then this estimated bias is subtracted from future GCM runs, assuming it remains constant over time. The most basic observation-driven procedure is the Delta method[1], where, by contrast, changes in mean temperature are estimated by comparing GCM future runs with those under historical forcing, and this trend is then added to the observational data. Both approaches implicitly assume that GCMs correctly capture changes in mean temperature, but they can result in temperature simulations that have very different higher-order characteristics.

An appealing property of observation-driven procedures like the Delta method is that they preserve attributes of the observations that are not explicitly accounted for in the simulation procedure, a property not shared by model-driven procedures. Figure 2.1, top row, provides a cartoon illustration of this difference between the two approaches. Here, the cartoon model predicts changes in mean but badly misrepresents the mean and covariance structure of the observations. In such a setting, simple bias correction will maintain the model's misrepresented covariance structure, while the Delta method yields a more realistic simulation (see Hawkins et al. (2013) for a less idealized example). More complicated versions

1. The meaning of the term "Delta method" in the geosciences, and in this work, is distinct from its typical use in the statistics literature for methods that employ Taylor expansions to derive asymptotic approximations of properties of functions of random variables.

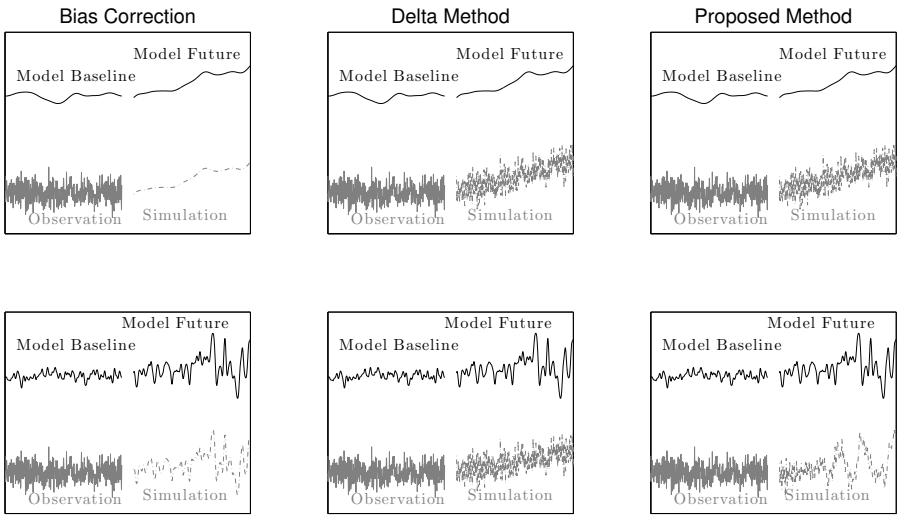

Figure 2.1: Cartoon illustration comparing strategies for simulating temperatures that combine information from a model and the observational record. Columns compare simple bias correction (left), the Delta method (center), and our proposed method (right). Top row, the model predicts changes in mean temperature but no changes in variability; in this case, our proposed method is equivalent to the Delta method. Bottom row, the model predicts changes in both mean and covariance. Simple bias correction does not retain the higher order properties of the observations, whereas the Delta method does not account for model changes in covariance; our proposed method does both.

of bias correction exist that attempt to correct for higher-order discrepancies between models and observations. Some correct for discrepancies in marginal distributions (e.g., Wood et al. (2004)), while others additionally attempt to correct rank correlation structures and inter-variable dependence structures (e.g., Piani and Haerter (2012); Vrac and Friederichs (2015)). While such methods are more sophisticated than the simple bias correction illustrated in Figure 2.1, they too will leave intact discrepancies between the model and observations not accounted for in the correction procedure. If impacts assessments require realistic simulations from the joint distribution of temperatures across space and time, our perspective is that this objective is more easily met by observation-driven methods.

Other routinely used simulation methods exist that do not fall as neatly within the model-driven/observation-driven dichotomy. For example, in simulations produced by stochastic weather generators (Semenov and Barrow, 1997; Wilks and Wilby, 1999), the observations

are replaced with synthetic data drawn from a stochastic model meant to mimic the distribution of the observations. The stochastic weather generator can then be modified to account for changes predicted by a climate model. The drawback of this approach is that a statistical model for the observations is required in addition to a model for GCM projected changes, whereas observation-driven methods only require the latter. Synthetically generated observations will be less realistic than the observations themselves. Other related methods in the statistics literature also attempt to use statistical models to blend information from observations with climate models to produce future simulations (e.g., Salazar et al. (2011)), but the proposed statistical models make very strong assumptions on the spatiotemporal distribution of the observations and the climate model realizations. Especially when projected changes from the historical climate are not very large, simulation methods should preserve features of the observed climate where possible. We therefore view observation-driven methods like the Delta method as likely to produce more realistic simulations than these methods as well.

We have advocated for observation-driven simulation methods, but an important limitation of the observation-driven Delta method described above is that it does not account for changes in variability (Figure 2.1, bottom row). Some extensions of the Delta method account for changes in marginal variance projected by a GCM (e.g., see again Hawkins et al. (2013)), but, again, since variability changes need not be uniform across all timescales of variation, changes in marginal variance are not a complete summary of GCM projected changes in variability. Leeds et al. (2015) introduced an extension of the Delta method that does account for changes in the full temporal covariance structure projected by a GCM, but their approach is applicable only for equilibrated climates, in which temperatures (after preprocessing for seasonality) can be assumed to be stationary in time. The Earth's climate, however, is and will continue to be in a transient state, in which temperatures will by definition be nonstationary in time. There is therefore an outstanding need for methods both to characterize changes in covariance in transient, nonstationary climates and to simulate temperatures in such climates.

In this work, we build on the work described in Leeds et al. (2015) to develop a methodology for generating observation-driven simulations of temperatures in future, transient climates that account for transient changes in both means and temporal covariances. In Figure 2.1, bottom row, our proposed method, unlike simple bias correction or the Delta method, both accounts for the relevant changes projected by the cartoon model and retains other distributional properties of the observations. Our method reduces to the Delta method in the case that the model predicts no changes in variability, and reduces to the method in Leeds et al. (2015) if the past and future climates are both in equilibrium. Since such a simulation uses projected changes in covariances from a GCM, our methodology must provide a way of modeling and estimating these changes in transient GCM runs. The transient, nonstationary setting adds substantial challenges, and so the statistical modeling of changes in covariance in transient GCM runs is a primary focus of this chapter.

As a final complication, since GCMs are extremely computationally intensive, it is not possible to run a GCM under every scenario relevant for impacts assessments. In the absence of a run for a scenario of interest, impacts modelers may instead rely on a GCM emulator, a simpler procedure that produces, for example, mean temperatures that mimic what the GCM would have produced had it been run. Our framework for simulations can use emulated rather than true GCM projections. For methods that emulate mean temperatures over forcing scenarios, see Castruccio et al. (2014) and references therein, and see the literature stemming from the pattern-scaling method of Santer et al. (1990). Much of the literature on climate model emulation has focused not on emulating model output across different forcing scenarios, but rather on emulating output with differing values of key climate model parameters (often for the purpose of selecting values of those parameters) (e.g., Chang et al. (2014); Rougier et al. (2009); Sansó et al. (2008); Sansó and Forest (2009); Sham Bhat et al. (2012); Williamson et al. (2013) and others). While the statistical concerns related to emulating climate models in parameter space are somewhat different from those of emulating in scenario space, the general idea remains the same: that one may use available climate

model runs to infer properties of a run that has not been produced. In our case, we require an emulator for the GCM changes in covariance in addition to a mean emulator. Our proposed statistical model can be used for this purpose, allowing our observation-driven procedure to simulate future temperatures in a potentially wide range of forcing scenarios.

Any simulation method that combines observations and model output (be it observation- or model-driven) involves the assumption that observations are less uncertain than is the output of GCMs. It is important to recognize that observational data products do themselves involve uncertainties that should be kept in mind when evaluating the appropriateness of a method for any particular use. In addition to intrinsic measurement uncertainty (e.g., calibration issues or time-varying changes), most datasets involve gridding data that are inhomogeneously distributed and come from different sources: raw observational data may be point-referenced (e.g., station data) or represent area averages (e.g., satellite measurements of radiances). Gridded data products are therefore affected by interpolation schemes and problems related to "spatial change of support" and misalignment (Gotway and Young, 2002), and care must be taken to appropriately combine raw observations to produce gridded products. To attempt to explicitly account for uncertainties, some observational data products are made available as ensembles (e.g., Morice et al. (2012)). In many cases (including this work) the observational data products used are "reanalyses," in which information from different sources are combined with a climate/weather model to produce a self-consistent solution. Reanalyses are therefore also affected by the model and data assimilation schemes used. For further discussion of uncertainty in the observational record, see IPCC (2013), Chapter 2, especially Box 2.1, and references therein. In this chapter, we do not explicitly account for uncertainties associated with the observational record but assume that, in practice, the use of any method that combines models and observations also involves the appropriate evaluation of observational data products relative to raw model output. One method of evaluation is to repeat any procedure used with multiple data sources.

The remainder of this article is organized as follows. In Section 2.2, we motivate and

describe our procedure for observation-driven simulations of future temperatures in transient climates that accounts for projected changes in both means and temporal covariances. In Section 2.3, we describe a GCM ensemble that we use to illustrate our methodology. In Section 2.4, we describe a statistical model for the changes in temporal covariances observed in this GCM ensemble that can be used as an emulator for these changes; we also discuss the estimation of this statistical model. In Section 2.5, we discuss results of applying our model to the GCM output and generating corresponding simulations, and evaluate the quality of our model in emulating the GCM projections. In Section 2.6, we give some concluding remarks and highlight areas for future research. Supplementary material may be found in the Appendix, Section 2.7.

2.2 Observation-driven simulations of temperatures in future transient climates

Our goal is to provide a simulation of future temperatures in a transient climate under a known forcing scenario. In light of the preceding discussion, this simulation should reflect knowledge of the changes in the mean and covariance structure of future temperatures under that scenario, but should otherwise preserve properties of the observed temperature record.

Our proposed procedure is motivated by an idealization of the problem, supposing that the future changes in mean and temporal covariance structure are known. Following this motivation, we describe some modifications to the proposed procedure that we argue make the procedure more useful in practical settings when changes in mean and covariance must be estimated from, for example, GCM runs.

2.2.1 Idealization

Consider a family of multivariate (i.e., spatially referenced) Gaussian time series, $z_l^{(s)}(t)$ at times $t = ..., -1, 0, 1, ...$ and locations $l = 1, ..., L$, indexed by $s \in \mathcal{S}$, some set of scenarios.

Write $\mu_l^{(s)}(t)$ for the unknown mean of $z_l^{(s)}(t)$ and assume that at each location, $z_l^{(s)}(t)$ has an unknown evolutionary spectrum, $a_l^{(s)}(t,\omega)$; for details on processes with evolutionary spectra, see Priestley (1981). Processes with temporally varying covariance structures in general have been discussed extensively in the literature. An overview of a theoretical framework for understanding locally stationary processes, closely tied to the Priestley model, can be found in Dahlhaus (2012) and the references therein. Our focus in this chapter is on spectral methods because we view evolutionary spectra to be an intuitive way to characterize time-varying covariances and because the process's corresponding spectral representation has useful implications for our simulation procedure. Most importantly, $z_l^{(s)}(t)$ has, at each location and for each s, the spectral representation

$$z_l^{(s)}(t) = \mu_l^{(s)}(t) + \int_{-\pi}^{\pi} e^{i\omega t}\sqrt{a_l^{(s)}(t,\omega)}\, d\xi_l^{(s)}(\omega),$$

where $\xi_l^{(s)}(\omega)$ is a mean zero process with orthogonal increments and unit variance; that is, $E[d\xi_l^{(s)}(\omega)d\xi_l^{(s)}(\omega)^*] = d\omega$ and $E[d\xi_l^{(s)}(\omega)d\xi_l^{(s)}(\omega')^*] = 0$ for $\omega \neq \omega'$, with * denoting the complex conjugate. Here and throughout this chapter, we restrict our attention to nonstationary processes with evolutionary spectral representations whose transfer functions, $\sqrt{a_l^{(s)}}$, are real and positive.

Suppose that we observe the time series under one scenario, $z_l^{(0)}(t)$, for times $t = 1, ..., N_0$. Call $z_l^{(0)}(t)$ the observed time series; given the observed time series, we would like to generate a simulation of the same length as the observations, but approximately equal in distribution to an unobserved time series, $z_l^{(s)}(t)$, for a given s. Since both $z_l^{(0)}(t)$ and $z_l^{(s)}(t)$ are Gaussian, there is a class of affine transformations of $z_l^{(0)}(t)$ that is equal in distribution to $z_l^{(s)}(t)$; indeed, writing $\Sigma_l^{(0)}$ and $\Sigma_l^{(s)}$ for the covariance matrices of the observed and unobserved time series at location l, we have that (marginally, at each location)

$$\mathbf{z}_l^{(s)} =_{\mathcal{D}} \boldsymbol{\mu}_l^{(s)} + (\Sigma_l^{(s)})^{1/2}(\Sigma_l^{(0)})^{-1/2}(\mathbf{z}_l^{(0)} - \boldsymbol{\mu}_l^{(0)})$$

for any matrix square root, where $\mathbf{x}$ denotes the vector with entries $x(t)$. While it is not immediately obvious that this fact is helpful, since the means and covariances of the two time series are unknown and at least for the unobserved time series cannot be directly estimated, we will describe a setting in which it is possible to compute this transformation (approximately) without fully knowing the means and covariances of the two time series.

The covariances of the observed and unobserved time series may be written as

$$\begin{aligned}(\Sigma_l^{(0)})_{t,t'} &= \int_{-\pi}^{\pi} e^{i\omega(t-t')}\sqrt{a_l^{(0)}(t,\omega)a_l^{(0)}(t',\omega)^*}\,d\omega, \text{ and} \\ (\Sigma_l^{(s)})_{t,t'} &= \int_{-\pi}^{\pi} e^{i\omega(t-t')}\sqrt{a_l^{(s)}(t,\omega)a_l^{(s)}(t',\omega)^*}\,d\omega,\end{aligned}$$

which follows immediately from the processes' spectral representations. Guinness and Stein (2013b) showed that, under some regularity conditions on the evolutionary spectra, these matrices can be approximated as

$$\begin{aligned}\Sigma_l^{(0)} &\approx C_{N_0}\left(\sqrt{a_l^{(0)}}\right)C_{N_0}\left(\sqrt{a_l^{(0)}}\right)^H, \text{ and} \\ \Sigma_l^{(s)} &\approx C_{N_0}\left(\sqrt{a_l^{(s)}}\right)C_{N_0}\left(\sqrt{a_l^{(s)}}\right)^H\end{aligned}$$

where H denotes the conjugate transpose and, generically for some function $A(t,\omega)$ in time and frequency, $C_N(A)$ is the $N \times N$ matrix with entries

$$C_N(A)_{t,j} = \sqrt{\frac{2\pi}{N}}\,A(t, 2\pi(j-1)/N)e^{2\pi i(j-1)t/N}$$

for $t, j = 1, ..., N$. (In the setting where A is constant in time, $C_N(A)$ is the inverse discrete time Fourier transform scaled by A and the result is the well-known result that the discrete time Fourier transform approximately diagonalizes the covariance matrix for a stationary time series observed at evenly spaced intervals.) The following transformation of $z_l^{(0)}(t)$ is

therefore, marginally at each location, l, approximately equal in distribution to $z_l^{(s)}(t)$:

$$\mathbf{z}_l^{(s,0)} = \boldsymbol{\mu}_l^{(s)} + C_{N_0}\left(\sqrt{a_l^{(s)}}\right) C_{N_0}\left(\sqrt{a_l^{(0)}}\right)^{-1} (\mathbf{z}_l^{(0)} - \boldsymbol{\mu}_l^{(0)}). \tag{2.1}$$

The crucial observation, however, is that (2.1) can be computed exactly without fully knowing the means and covariances of the observed and unobserved time series. Indeed, suppose that what we are given are not the means and evolutionary spectra of the processes themselves, but some other substitute set of functions $\tilde{\mu}_l^{(s)}(t)$ and $\tilde{a}_l^{(s)}(t,\omega)$ satisfying, for each scenario s and at each location l,

$$\mu_l^{(s)}(t) - \tilde{\mu}_l^{(s)}(t) = c_l, \text{ and } \frac{a_l^{(s)}(t,\omega)}{\tilde{a}_l^{(s)}(t,\omega)} = k_l(\omega), \tag{2.2}$$

for some unknown constant c_l and some unknown function $k_l(\omega)$ that is constant in time. This situation is analogous to our actual predicament, where GCM runs are assumed to be more informative about changes than absolute levels; indeed, one consequence of the assumptions (2.2) is that the substitute means and evolutionary spectra change in the same way as their true counterparts, so, for instance, we may write

$$\Delta_l^{(s,0)}(t) \equiv \mu_l^{(s)}(t) - \mu_l^{(0)}(t) = \tilde{\mu}_l^{(s)}(t) - \tilde{\mu}_l^{(0)}(t), \text{ and}$$

$$\rho_l^{(s,0)}(t,\omega) \equiv \frac{a_l^{(s)}(t,\omega)}{a_l^{(0)}(t,\omega)} = \frac{\tilde{a}_l^{(s)}(t,\omega)}{\tilde{a}_l^{(0)}(t,\omega)},$$

for the known changes in means and covariance structures. The assumption that GCM mean temperatures are off by a constant compared to real temperatures is essentially the assumption underlying both simple bias correction and the Delta method as described in Section 2.1, and we view the assumption on the evolutionary spectra as a natural extension to covariances; all existing simulation methods that we are aware of implicitly or explicitly make the same or similar assumptions (except the simple Delta method, which assumes no

changes in variability at all).

Under these assumptions, (2.1) may be rewritten as

$$\mathbf{z}_l^{(s,0)} = \boldsymbol{\mu}_l^{(0)} + \boldsymbol{\Delta}_l^{(s,0)} + C_{N_0}\left(\sqrt{\tilde{a}_l^{(s)}}\right) C_{N_0}\left(\sqrt{\tilde{a}_l^{(0)}}\right)^{-1} (\mathbf{z}_l^{(0)} - \boldsymbol{\mu}_l^{(0)}), \tag{2.3}$$

where, to reiterate, (2.1) and (2.3) are equal under the assumptions (2.2) because

$$C_{N_0}\left(\sqrt{\tilde{a}_l^{(s)}}\right) = C_{N_0}\left(\sqrt{a_l^{(s)}}\right) \mathrm{diag}(1/k_l),$$

where $\mathrm{diag}(1/k_l)$ is the diagonal matrix with entries $1/k_l(\omega_j)$, so

$$C_{N_0}\left(\sqrt{\tilde{a}_l^{(s)}}\right) C_{N_0}\left(\sqrt{\tilde{a}_l^{(0)}}\right)^{-1} = C_{N_0}\left(\sqrt{a_l^{(s)}}\right) C_{N_0}\left(\sqrt{a_l^{(0)}}\right)^{-1}.$$

In light of (2.3), our proposed simulation can be computed as long as one knows (or, more realistically, can estimate) just the mean of the observed time series as well as the substitute evolutionary spectra and changes in mean. The procedure described by (2.3) is what is illustrated in the bottom right panel of Figure 2.1.

In the case that there are no changes in covariance structure, so $a_l^{(s)}(t,\omega) = a_l^{(0)}(t,\omega)$ and the same for their substitute counterparts, the simulation procedure (2.3) is equivalent to the Delta method as described in Section 2.1. In the case that both $a_l^{(s)}(t,\omega)$ and $a_l^{(0)}(t,\omega)$ are constant in time, so the de-meaned time series are stationary, the procedure is the same as that described in Leeds et al. (2015). Our proposal is therefore a generalization of those two procedures, describing an observation-driven simulation that transforms one observed time series (possibly itself from a transient climate) to a simulation under a new (future, transient) scenario.

2.2.2 Practical modifications to idealized procedure

In practice, we do not actually know even substitute versions of the future changes in mean and covariance structure, so the procedure we have described in the preceding section must be modified to be made useful.

A key assumption underlying our methodology is that GCM runs are informative about at least some aspect of the changes in mean and covariance structure of the real temperatures; however, it need not be true that the assumptions (2.2) will be satisfied by taking $\mu_l^{(s)}(t)$ and $a_l^{(s)}(t,\omega)$ to be the means and evolutionary spectra corresponding to real temperatures and taking $\tilde{\mu}_l^{(s)}(t)$ and $\tilde{a}_l^{(s)}(t,\omega)$ to be those corresponding to temperatures under GCM runs. One possible objection to these assumptions is that both the observations and the GCM runs will exhibit nonstationarity in mean and variance due to seasonality, and it is at least plausible that the GCM representation of these seasonal cycles will differ from that of the observations. Leeds et al. (2015) argued that the seasonality can be reasonably represented as a uniformly modulated process (see Priestley (1981)) plus a mean seasonal cycle. That is, writing $T_l^{(s)}(t)$ for the true temperatures at time t and location l in scenario s, and d for the day of the year, we assume that

$$T_l^{(s)}(t) = \mu_l^{(s)}(t) + m_l^{(T)}(d) + D_l^{(T)}(d)(z_l^{(s)}(t) - \mu_l^{(s)}(t))$$

where $m_l^{(T)}(d)$ and $D_l^{(T)}(d)$ represent seasonal cycles in means and marginal variances, and $z_l^{(s)}(t)$ has mean $\mu_l^{(s)}(t)$ and evolutionary spectrum $a^{(s)}(t,\omega)$ as above; assume a similar form for the GCM runs. (In the following, we will allow $\mu_l^{(s)}(t)$ to reflect changes in the mean seasonal cycle from $m_l^{(T)}(d)$ but for simplicity will assume that $a^{(s)}(t,\omega)$ has no seasonal structure; see Section 2.3.1 for details.) We estimate the seasonal cycles in mean and variability in the observations according to the methods described in Leeds et al. (2015); the mean seasonal cycle is modeled with the first ten annual harmonics and estimated via least squares, whereas the annual cycle in variability is estimated by a normalized moving average

of windowed variances that has been averaged across years. We will assume that (2.2) is reasonable taking $\mu_l^{(s)}(t)$, $a^{(s)}(t,\omega)$, $\tilde{\mu}_l^{(s)}(t)$, and $\tilde{a}^{(s)}(t,\omega)$ to be the means and evolutionary spectra of the deseasonalized components of real and GCM temperatures.

Even assuming that (2.2) holds for the deseasonalized components of observed and GCM temperatures, what we are given are the real and GCM temperatures themselves, not their means and evolutionary spectra. As such, the quantities necessary to compute (2.3) must be estimated using the available data. While it would be possible to estimate the corresponding evolutionary spectra from the GCM runs directly, note that (2.3) works for any substitute function in time and frequency that, for each frequency, is proportional to the true evolutionary spectra at all times. If such a function must be estimated, there is presumably a statistical advantage to estimating a function that is relatively flat across frequencies. In GCM experiments, it is fairly typical to have a control run under an equilibrated (often preindustrial) climate, in which at least the deseasonalized temperatures can be viewed as a stationary process. Writing $s = B$ for this equilibrated scenario, and $\tilde{a}_l^{(B)}(\omega) \equiv \tilde{a}_l^{(B)}(t,\omega)$ for the corresponding spectral density of the deseasonalized component of the equilibrated GCM temperatures, then if (2.2) holds, it will also be true that we can write

$$\rho_l^{(s,B)}(t,\omega) \equiv \frac{a^{(s)}(t,\omega)}{a^{(B)}(\omega)} = \frac{\tilde{a}^{(s)}(t,\omega)}{\tilde{a}^{(B)}(\omega)},$$

in which case $a^{(s)}(t,\omega)/\rho_l^{(s,B)}(t,\omega) = a^{(B)}(\omega)$ and (2.2) still holds if one replaces $\tilde{a}^{(s)}(t,\omega)$ with $\rho_l^{(s,B)}(t,\omega)$. Moreover, we expect that the functions $\rho_l^{(s,B)}(t,\omega)$ will be much flatter than the functions $\tilde{a}_l^{(s)}(t,\omega)$ over the range of scenarios considered reasonable, so we expect that there should be some advantage in estimating these ratios rather than the evolutionary spectra themselves from the GCM runs.

Writing $T_l^{(s,0)}(t)$ for our simulation of the true temperatures under scenario s, our pro-

posed procedure is therefore

$$\begin{aligned} \mathbf{T}_l^{(s,0)} &= \hat{\boldsymbol{\mu}}_l^{(0)} + \hat{\mathbf{m}}_l^{(T)} + \hat{\boldsymbol{\Delta}}_l^{(s,0)} \\ &\quad + \mathrm{diag}(\hat{\mathbf{D}}_l^{(T)}) C_{N_0} \left(\sqrt{\hat{\rho}^{(s,B)}} \right) C_{N_0}^{-1} \left(\sqrt{\hat{\rho}^{(0,B)}} \right) (\hat{\mathbf{z}}_l^{(0)} - \hat{\boldsymbol{\mu}}_l^{(0)}), \end{aligned} \tag{2.4}$$

where $\hat{x}$ generically represents an estimate of the quantity x. In words, the procedure is, in order, (i) estimate and remove seasonality and mean trend in the observational record; (ii) estimate the future changes in mean and marginal spectra using an ensemble of GCM runs; (iii) decorrelate the detrended and deseasonalized observations using the estimated substitute function, $\hat{\rho}^{(0,B)}$, obtained in step (ii); (iv) apply the estimated changes in spectra, $\hat{\rho}^{(s,B)}$, to the decorrelated data and invert the transformation in (iii); and (v) replace the seasonal cycles in means and variability, and add the new changes in mean.

The practicality of (2.4) depends both on our ability to obtain good estimates of all of the involved quantities and on our ability to compute the simulation efficiently. Castruccio et al. (2014) and Leeds et al. (2015) collectively describe methods that essentially can be used to estimate all of the necessary quantities in the procedure except, crucially, the changes in evolutionary spectra. In the following, we will discuss modeling and estimating these functions from an ensemble of GCM runs. The particular statistical model we develop allows for efficient computation of the simulation.

2.3 Description of GCM ensemble

We study changes in the distribution of daily temperatures in an ensemble of GCM runs made with the Community Climate System Model Version 3 (CCSM3) (Collins et al., 2006; Yeager et al., 2006) at T31 atmospheric resolution (a 48×96 grid with a resolution of approximately $3.75^\circ \times 3.75^\circ$) and a 3° resolution for oceans. All runs require a long spin up; the realizations in our ensemble are initialized successively at ten-year intervals of the NCAR b30.048 preindustrial control run. Each transient realization is then forced by historical CO_2

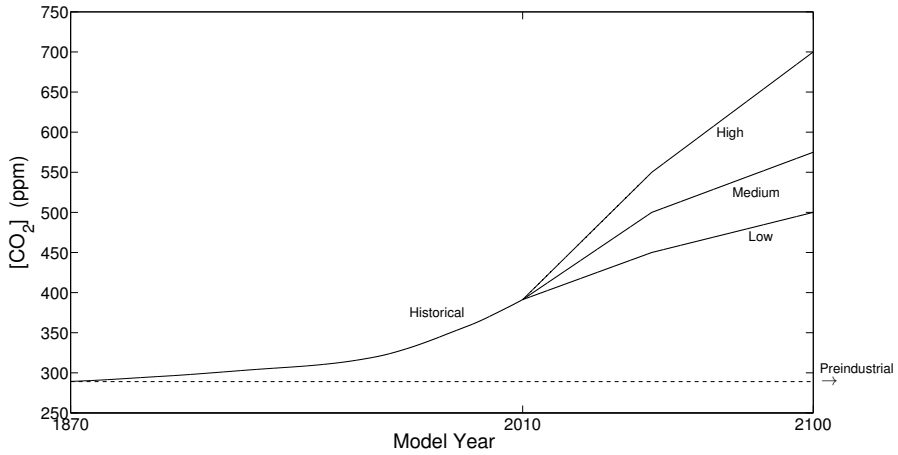

Figure 2.2: GCM ensemble [CO_2] trajectories. The historical scenario begins in 1870 and follows observed [CO_2] until 2010, after which it branches into the three future scenarios increasing at different rates until 2100. The preindustrial run maintains 1870-level [CO_2] until the year 4600, but we use only the last 1,500 years of that run. The ensemble includes eight realizations each under the historical, high, and low scenarios, five under the medium scenario, and one under the preindustrial scenario.

concentrations (herein, [CO_2]) from years 1870-2010, at which point the ensemble branches into three future increasing [CO_2] scenarios for the years 2010-2100, which we name the "high," "medium," and "low" concentration scenarios (Figure 2.2). For each scenario, we have a modest number of realizations (eight realizations from the historical, high, and low scenarios, and five realizations from the medium scenario), so the transient ensemble consists of about 1.1 million observations at each grid cell, or about 5 billion observations in total. As is typical, we will assume throughout that the ensemble members can be viewed as statistically independent realizations due to the system's sensitivity to initial conditions.

The focus of our investigation is on changes in temporal covariance structure in transient (nonstationary) runs of the GCM, but for the reasons described in Section 2.2.2, it is helpful to have a representation of the model's climate in a baseline, equilibrated state. For this purpose, we use a single, long run forced under preindustrial [CO_2] (289 ppm) for an additional 2,800 years past the control run initialization to ensure that the run is fully equilibrated, from which we use the last 1,500 years, or about 0.5 million days, for a total

of about 2.5 billion observations under preindustrial [CO_2].

In the following, we will index the members of the GCM ensemble by their [CO_2] scenario, $s \in \{B, 0, H, M, L\}$, denoting, respectively, the baseline preindustrial, historical, high, medium, and low scenarios.

2.3.1 Data preprocessing

The primary inferential aim of this work is to obtain estimates of $\rho_l^{(s,B)}(t, \omega)$, the changes in marginal evolutionary spectra of the deseasonalized component of daily temperatures in the GCM under scenario s compared to the preindustrial climate. The GCM runs have, accordingly, been preprocessed to remove means and seasonal cycles of variability.

As with the observed temperatures in Section 2.2, we represent temperatures in the preindustrial run at each grid cell as a uniformly modulated process plus a mean seasonal cycle and retain the stationary component of this process. That is, write $y_l^{(B)}(t)$ for the temperature in the raw, equilibrated preindustrial GCM run at time t and location l, and again write d for the day of the year (the GCM does not have leap years, so $d \in \{1, ..., 365\}$). We represent these as

$$y_l^{(B)}(t) = \tilde{m}_l(d) + \tilde{D}_l(d) x_l^{(B)}(t),$$

where $\tilde{m}_l(d)$ and $\tilde{D}_l(d)$ are the estimated seasonal cycles in mean and marginal variance, and $x_l^{(B)}(t)$ is assumed to be stationary in time. The mean seasonal cycle and the seasonal cycle of marginal variance are estimated as described in Leeds et al. (2015), as also in the preceding section for the observational data. We retain $x_l^{(B)}(t)$, the deseasonalized component.

Temperatures in the transient runs, on the other hand, will in general have an evolving mean in addition to nonstationarity due to seasonality. Write $y_{l,r}^{(s)}(t)$ for the temperature in the r'th realization of the transient scenario s at time t and location l, and assume the representation

$$y_{l,r}^{(s)}(t) = \tilde{m}_l(d) + \tilde{\mu}_l^{(s)}(t) + \tilde{D}_l(d) x_{l,r}^{(s)}(t),$$

where $\tilde{\mu}_l^{(s)}(t)$ represents an estimate of the evolving mean under scenario s (possibly including changes in the mean seasonal cycle) and $x_{l,r}^{(s)}(t)$ is assumed to be some mean zero, but nonstationary, process. For simplicity we assume that the seasonal cycles of marginal variability do not evolve in time (overall marginal variability is still allowed to change in $x_{l,r}^{(s)}(t)$, but such changes are assumed the same for each season). While the change in mean is needed for the simulation (2.4) (see Appendix 2.7.1 for details on its estimation), we would like to work with mean zero processes to estimate the changes in covariance structure. We therefore first remove from each transient realization the scenario average:

$$\check{y}_{l,r}^{(s)}(t) = y_{l,r}^{(s)}(t) - \frac{1}{R_s}\sum_{k=1}^{R_s} y_{l,k}^{(s)}(t),$$

where R_s is the number of realizations in the ensemble under scenario s. The resulting contrasts, $\check{y}_{l,r}^{(s)}(d)$, have mean zero, but still exhibit seasonal cycles in marginal variance. We thus retain the deseasonalized contrasts

$$q_{l,r}^{(s)}(t) = \frac{\check{y}_{l,r}^{(s)}(t)}{\tilde{D}_l(d)}. \tag{2.5}$$

While we view each run, $y_{l,r}^{(s)}(t)$, as independent, the contrasts, $q_{l,r}^{(s)}(t)$, are of course not independent across realizations within a given scenario.

We assume that the deseasonalized component of the preindustrial run, $x_l^{(B)}(t)$, has unknown marginal spectral density $\tilde{a}_l^{(B)}(\omega)$, and the deseasonalized component of the transient runs, $x_{l,r}^{(s)}(t)$, has unknown evolutionary spectrum $\tilde{a}_l^{(s)}(t,\omega)$. While $x_l^{(B)}(t)$ and $x_{l,r}^{(s)}(t)$ will not be equal in distribution to the (deseasonalized) real-world temperatures, past or future, we assume that the true changes in evolutionary spectra under scenario s are equal to those of the GCM, so the deseasonalized GCM and observed temperatures satisfy (2.2) and, in particular,

$$\rho_l^{(s,B)}(t,\omega) \equiv \frac{a_l^{(s)}(t,\omega)}{a_l^{(B)}(\omega)} = \frac{\tilde{a}_l^{(s)}(t,\omega)}{\tilde{a}_l^{(B)}(\omega)}.$$

In the following section, we discuss modeling and estimating these changes in evolutionary spectra.

2.4 GCM projected changes in temporal covariance

We describe a methodology for modeling and estimating the changes in covariance structure in a GCM as a function of a $[CO_2]$ scenario. Our goal is not only to describe the changes in covariance in scenarios within our ensemble, but also to provide an emulator for the GCM changes in covariance for scenarios for which we have no runs. To the extent that the model we propose describes the GCM changes across the range of $[CO_2]$ scenarios in our ensemble, the resulting emulator may be expected to provide good predictions of the GCM changes, at least for scenarios in some sense within the range spanned by our ensemble.

2.4.1 A model for GCM changes in temporal covariance

An important insight stated in Castruccio et al. (2014) is that changes in the distribution of temperatures in transient GCM runs under a $[CO_2]$ forcing scenario should be describable in terms of the past trajectory of $[CO_2]$. More specifically, writing $[CO_2](t)$ for the CO_2 concentration at time t, the distribution of temperature at time t is determined by the function

$$f(t') = [CO_2](t - t'), \text{ for } t' > 0,$$

where f does not depend on t, so one does not need a different emulator for every time t. Providing useful statistical emulators for changes in the distribution of temperatures in transient GCM runs then depends on our ability to find useful functionals of the past $[CO_2]$ trajectory that help explain those changes.

One potentially useful summary of the past trajectory of $[CO_2]$ for a given scenario is in fact the change in regional mean temperature relative to the preindustrial climate. We denote this change as $\bar{\Delta}_S^{(s,B)}(t)$ for region S. In this work, we have subdivided the T31 grid

into the same 47 regions as in Castruccio et al. (2014), chosen to be relatively homogeneous but still large enough to substantially reduce inter-annual variations. We estimate $\bar{\Delta}_S^{(s,B)}(t)$ in each region using a modification of the mean emulator described there (see Appendix 2.7.1).

While the changes in regional mean temperature are themselves useful summaries of the past trajectory of $[CO_2]$, it need not be true that temporal covariance structures will be the same if $\bar{\Delta}_S^{(s,B)}(t) = \bar{\Delta}_S^{(s',B)}(t')$ for two scenarios s and s' at two different time points t and t'. In particular, the rate of change of the evolution of regional mean temperatures ($\partial_t \bar{\Delta}_S^{(s,B)}(t)$ for scenario s) may capture some additional aspect of the changing climate that is also relevant for explaining changes in covariances.

We have indeed found that the following model usefully describes the changes in temporal covariances in scenarios in our ensemble:

$$\log \rho_l^{(s,B)}(t,\omega) = \delta_{l0}(\omega)\bar{\Delta}_S^{(s,B)}(t) + \delta_{l1}(\omega)\partial_t \bar{\Delta}_S^{(s,B)}(t). \tag{2.6}$$

In the case that $\delta_{l0}(\omega)$ and $\delta_{l1}(\omega)$ are constant functions, for example, (2.6) describes a uniformly modulated process. More generally, $\delta_{l0}(\omega)$ and $\delta_{l1}(\omega)$ describe the patterns of changes in variability across frequencies associated with changes in regional mean temperature and the rate of change of regional mean temperature, respectively.

Since each $\delta_{li}(\omega)$ is not scenario-dependent, model (2.6) can be thought of as an emulator for the GCM changes in covariance structure. That is, given an emulator for the regional mean temperature changes in the scenario of interest, (2.6) provides a prediction of the GCM changes in covariance structure under that scenario. In Section 2.5, we will discuss how well this model describes changes across the scenarios in our ensemble. Note that a model like (2.6) is unlikely to hold generically for all $[CO_2]$ scenarios. In particular, such a model would be unlikely to fully capture the changes in variability in scenarios where $[CO_2]$ changes instantaneously; such scenarios are typically not considered realistic. Addi-

tionally, since the changes in covariance in (2.6) depend on the $[CO_2]$ trajectory through the corresponding *changes* in mean, this model will not fully capture changes in variability that depend on *absolute* temperatures through, for example, phase changes between ice and water (see Figure 2.5 and discussion in Section 2.5.2). The model will also not fully capture GCM behavior if that behavior involved abrupt changes in the distribution of temperatures even under relatively smooth forcing scenarios (nonlinear responses to forcing); however, the GCM we study does not exhibit such behavior over the range of $[CO_2]$ scenarios we study. We have found that the model is a useful description of changes in variability in scenarios like those in the ensemble we use here, where $[CO_2]$ changes slowly and relatively smoothly over time and in locations not involving changing ice margins over the course of the scenario.

Estimating $\delta_{li}(\omega)$

To estimate the functions $\delta_{l0}(\omega)$ and $\delta_{l1}(\omega)$, we adopt the intuitive approach for likelihoods for processes with evolutionary spectra where the usual periodogram in the Whittle likelihood is replaced with local periodograms over smaller blocks of time (Dahlhaus, 1997). While we view the local Whittle likelihood approach as most suitable in our setting, several other alternative methods for estimating evolutionary spectra have been proposed. Neumann and Von Sachs (1997) used a wavelet basis expansion; Ombao et al. (2002) used smooth, localized complex exponential basis functions; Dahlhaus (2000) proposed another likelihood approximation that replaces the local periodogram in the earlier work with the so-called pre-periodogram introduced by Neumann and Von Sachs (1997); Guinness and Stein (2013b) provided an alternative generalization of the Whittle likelihood that they argued, at least in the settings they studied, is more accurate than the approximations given by either Dahlhaus (1997) or Dahlhaus (2000). However, an advantage of the local Whittle likelihood approach is that in addition to being intuitive, the corresponding score equations are computationally easier to solve in our setting when the evolutionary spectra evolve very slowly in time so the local periodograms can be taken over large blocks of time. Computation is an especially

important consideration when estimating a semiparametric model such as (2.6). Furthermore, the results from Guinness and Stein (2013b) suggest that the local Whittle likelihood approach may yield point estimates that are close to optimal even when the likelihood approximation itself is inaccurate, whereas, for example, they demonstrated that the approach based on the pre-periodogram can give unstable estimates.

In this work, we interpret the local Whittle likelihood approach as follows. We divide each contrast time series, $q_{l,r}^{(s)}(t)$, defined in (2.5) and of length N_s, into blocks of length M (for simplicity take M to be a common factor of each N_s). In our setting, we take $M = 10$ years, but since temperature variability changes very slowly over time in the scenarios we analyze, the conclusions are not very sensitive to the choice of M; the results are essentially the same taking $M = 5$ years or $M = 30$ years, for example. Upon choosing M, then for the time block, location, realization, and scenario indexed by b, l, r, and s, respectively, define the local periodogram of the contrast time series at frequencies $\omega_j = 0, 2\pi/M, ..., 2\pi$ as

$$I_{b,l,r}^{(s)}(\omega_j) = \frac{1}{2\pi M}\left|\sum_{t=1}^{M} q_{l,r}^{(s)}(t + M(b-1))e^{-it\omega_j}\right|^2. \tag{2.7}$$

It is straightforward to show that the Whittle likelihood for each time block, location, and scenario depends on each $I_{b,l,r}^{(s)}(\omega_j)$ only through the average across realizations,

$$\bar{I}_{b,l}^{(s)}(\omega_j) = \frac{1}{R_s}\sum_{r=1}^{R_s} I_{b,l,r}^{(s)}(\omega_j). \tag{2.8}$$

Likewise, for the deseasonalized preindustrial run, define its periodogram as

$$I_l^{(B)}(\omega_j) = \frac{1}{2\pi N_B}\left|\sum_{t=1}^{N_B} x_l^{(B)}(t)e^{-it\omega_j}\right|^2. \tag{2.9}$$

In our setting, since $M < N_B$, $I_{b,l,r}^{(s)}(\omega_j)$ is defined on a coarser frequency scale than is $I_l^{(B)}(\omega_j)$, so for the purposes of estimating changes in spectra, it may be natural to aggregate

the baseline periodogram to the coarser scale; that is, write

$$\bar{I}_l^{(B)}(\omega_j) = \frac{M}{N_B} \sum_{k:\, -\frac{N_B}{2M} \le k < \frac{N_B}{2M}} I_l^{(B)}\left(\frac{2\pi(j+k)}{N_B}\right). \tag{2.10}$$

An approximate likelihood under model (2.6), marginally at each location l, may then be written as the sum of the local Whittle likelihoods of the transient runs and the Whittle likelihood corresponding to the aggregated periodogram of the baseline run (under the usual approximation that the periodogram ordinates are independent at distinct Fourier frequencies):

$$\begin{aligned}\mathcal{L}_l(\theta) = -\frac{1}{2}\sum_{s,b,j}\Big\{&(R_s-1)(\log\tilde{a}_l^{(B)}(\omega_j) + \bar{\Delta}_b^{(s,B)}\delta_{l0}(\omega_j) + \partial_t\bar{\Delta}_b^{(s,B)}\delta_{l1}(\omega_j)) \\ &+ R_s\bar{I}_{b,l}^{(s)}(\omega_j)e^{-(\log\tilde{a}_l^{(B)}(\omega_j)+\bar{\Delta}_b^{(s,B)}\delta_{l0}(\omega_j)+\partial_t\bar{\Delta}_b^{(s,B)}\delta_{l1}(\omega_j))}\Big\} \\ &- \frac{M}{2}\sum_j\left\{\log\tilde{a}_l^{(B)}(\omega_j) + \bar{I}_l^{(B)}(\omega_j)/\tilde{a}_l^{(B)}(\omega_j)\right\}.\end{aligned} \tag{2.11}$$

where $\theta = (\tilde{a}^{(B)}, \delta_{l0}, \delta_{l1})$ and where $\bar{\Delta}_b^{(s,B)}$ and $\partial_t\bar{\Delta}_b^{(s,B)}$ correspond to the values of $\bar{\Delta}^{(s,B)}(t)$ and $\partial_t\bar{\Delta}^{(s,B)}(t)$ for t at the midpoint of time block b. Here the (R_s-1) factor multiplying the log-determinant approximation takes into account that the contrasts, $q_{l,r}^{(s)}(t)$, are obtained by subtracting off the scenario average across realizations; see Castruccio and Stein (2013) for details.

The estimator maximizing (2.11), say,

$$\theta_l^* = \arg\max_\theta \mathcal{L}_l(\theta), \tag{2.12}$$

will yield very rough estimates of the functions $\tilde{a}_l^{(B)}(\omega)$, $\delta_{l0}(\omega)$, and $\delta_{l1}(\omega)$ because no smoothness has been enforced across frequencies. The baseline spectrum, $\tilde{a}_l^{(B)}(\omega)$, is not of particular interest to us, as this function is not required for the simulation (2.4). On the

other hand, maximizing (2.11) is clearly inadequate for estimating the functions of interest, $\delta_{l0}(\omega)$ and $\delta_{l1}(\omega)$.

A common approach for nonparametrically estimating the spectral density of a stationary process is to smooth its periodogram either by kernel methods or by penalized likelihood methods. For estimating ratios of spectra between two stationary processes, Leeds et al. (2015) adopted a penalized likelihood approach whereby the penalty enforced smoothness in the ratio. Here, we opt to smooth the rough estimates, δ^*_{l0} and δ^*_{l1}, using kernel methods; that is, for $i = 0, 1$ write as the final estimate for δ_{li}

$$\hat{\delta}_{li}(\omega_j) = \sum_k w_{k,j,i} \delta^*_{li}(\omega_{j+k}), \tag{2.13}$$

where $w_{k,j,i}$ are weights (possibly varying with j and i) satisfying $\sum_k w_{k,j,i} = 1$ for each i and j. In practice, we use weights corresponding to a kernel with a variable bandwidth that is allowed to decrease at lower frequencies. The reason for the variable bandwidth is that in the GCM runs we have analyzed, we have observed that the log ratio of spectra is typically less smooth at very low frequencies compared to higher frequencies. For details on the form of the weights and the bandwidth selection procedure we use to choose them, see Appendix 2.7.1. While the penalized likelihood approach described in Leeds et al. (2015) may be adapted for this setting, we view the kernel smoothing approach as more straightforward, especially when allowing for variable bandwidths, and have found that the approaches yield similar estimates when the bandwidth of the kernel is constant.

Approximate pointwise standard errors for each $\hat{\delta}_{li}(\omega_j)$, and for the corresponding estimate of $\log \rho_l^{(s,B)}(t, \omega_j)$, may also be computed; these are described in Appendix 2.7.3. Having estimated our model, we need to compute the observation-driven simulations. Computing (2.4) efficiently is important; this is described in Appendix 2.7.4.

2.5 Results

In this section, we show results from applying the procedures described in the previous sections. We first estimate the model described in Section 2.4 using our ensemble of CCSM3 runs described in Section 2.3. We then investigate the success of our proposed model in describing the changes in covariances in our GCM ensemble, evaluate the quality of our model when used as an emulator, and describe insights into the climate system that our estimated model provides. As an illustration, we use an observational data product to build a simulation of future temperatures via the methods outlined in Section 2.2.

In the simulation, we use temperatures from the NCEP-DOE Climate Forecast System Reanalysis (CFSR) (Saha et al., 2010) as a surrogate for observational data. CFSR is the latest version of the global reanalysis produced by the National Centers for Environmental Prediction (NCEP). The reanalysis is run at T62 resolution (about a 1.875° grid); we regrid to T31 to match the resolution of our model output, using an area-conserving remapping scheme. Validation studies suggest that modern reanalyses, including CFSR, have relatively small mean biases and variability discrepancies relative to station observations (Decker et al., 2012), and CFSR appears, for example, to better represent inter-annual variability than older NCEP reanalyses (Wang et al., 2011). Nevertheless, as noted in Section 2.1, since there are differences between observational products and since our method does not account for uncertainties in the observations themselves, it is advisable for users to explore the effect that the choice of data product has on the resulting simulation.

2.5.1 Model changes in variability

In CCSM3, changes in variability in evolving climates can be primarily characterized using changes in mean temperature, with a smaller contribution by the rate of change of warming (corresponding to terms δ_{l0} and δ_{l1} in (2.6), respectively). As a consequence, the projected patterns of changes in variability at a given time in a given future scenario largely correspond

to the patterns observed in Leeds et al. (2015): the GCM projects decreases in short timescale variability at most locations, but increases in longer timescale variability in some regions, especially at lower latitudes (Figure 2.3, left).

The differences in variability between scenarios due to different rates of warming are small compared to the overall projected changes in variability, but also exhibit patterns. To illustrate, we compare changes in variability under the low scenario at year 2100 to the corresponding changes under the high scenario in the year of that scenario experiencing the same regional mean temperatures as at 2100 in the low scenario (Figure 2.3, right); this year varies by region, ranging from 2037 to 2044. An analogous figure is given in Appendix 2.7.5 that shows the estimated changes in variability in each of the three scenarios at years corresponding to the same change in regional mean temperature. In about 75% of all locations, and especially in mid- and high-latitude ocean locations, the changes in variability under the high scenario are larger than under the low scenario. Larger changes under the high scenario than under the low scenario are an indication that variability changes are projected to be larger in a transient warming climate than in an equilibrated climate at the same temperature.

To illustrate how these changes in covariance structure are used in our proposed simulation, we simulate temperatures under the high scenario at a single grid cell in the Midwestern United States (Figure 2.4, which shows the observations in 2009-2010, our simulation 89 years in the future, and output from one of the GCM runs in the same timeframes). Mean temperatures warm, more strongly in the winter than in the summer at this location, and temporal variability decreases overall. More specifically, variability is projected to modestly decrease at higher frequencies and slightly increase at lower frequencies. At low frequencies, the projected log ratios are within two standard errors of zero, but at high frequencies are significantly smaller than zero. The extent to which such changes are important will of course depend on the impact domain of interest, and we have not carried out studies on specific events of potential interest (for example, hot spells).

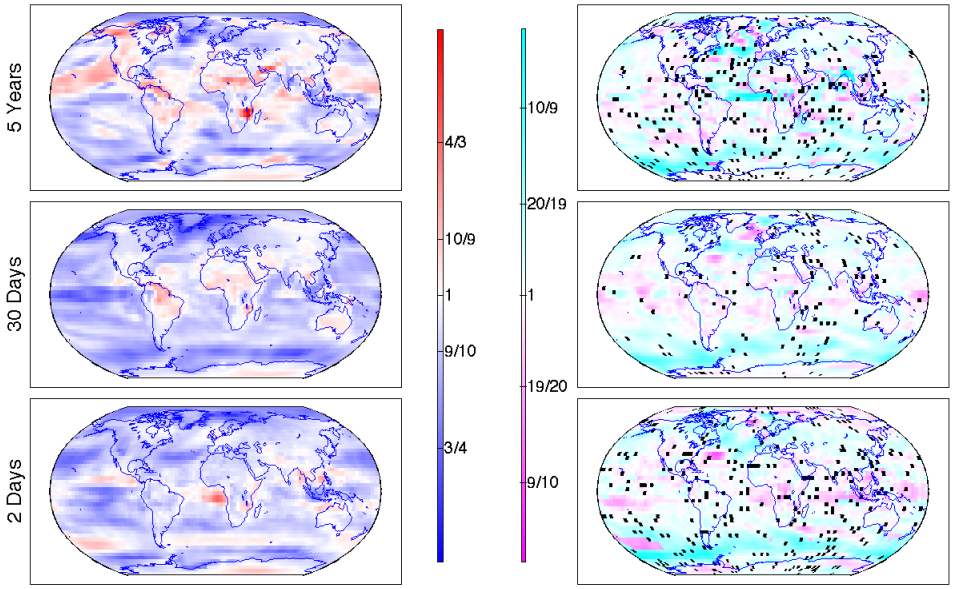

Figure 2.3: Left, estimates of changes in marginal spectra, at three frequencies, for the low scenario at model year 2100 compared to the preindustrial climate (i.e., $\rho_l^{(L,B)}(t,\omega)$ at year 2100 and the specified periods). Red indicates an increase in variability and blue a decrease in variability. Right, estimates the ratio of variability changes in the high to the low scenarios, for the year in the high scenario corresponding to the same change in regional mean temperature as at year 2100 in the low scenario (or the inverse of that ratio if the corresponding variability changes at the time, frequency, and location of interest are negative). Magenta grid cells indicate smaller changes in variability under the high scenario at the same temperature, whereas the cyan grid cells indicate bigger changes in variability. (Black grid cells indicate the roughly 5% of locations where the two estimates differ in sign, so comparing the relative magnitude of changes is not meaningful). Figure 2.10 repeats the left column maps for all the low, medium, and high scenarios.

The distribution of temperatures in the GCM differs strongly from that in the observations, with differences evident by eye in the raw time series and corresponding marginal densities. For example, the GCM has a stronger seasonal cycle than the observations and simulation, and greater variability in the winter months. See Appendix 2.7.5 for additional comparisons of the space-time covariance structures of the observations and the GCM runs: typically, we have found that temperatures in nearby grid cells are more coherent in the GCM than in the observations, and that the coherences do not change much between the historical period and the end of the high scenario (Figures 2.11-2.14). Our simulation procedure does not change the coherence structure of the observations. Collectively, this forms

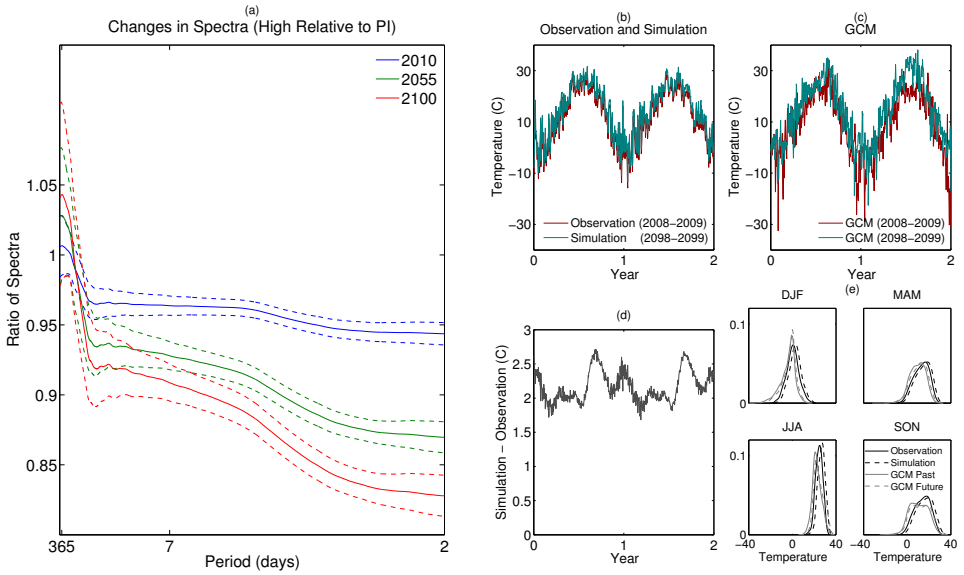

Figure 2.4: (a) estimates of $\rho_l^{(H,B)}(t,\omega)$ for years 2010, 2055, and 2100, at a Midwestern United States grid cell (dashed curves at $\pm$ two standard errors on the log scale). (b) part of the corresponding simulation computed by transforming the observational data at this grid cell via (2.4); the simulation was computed for the whole length of the observational data, but only the last two years are displayed. (c) a run of the GCM in the years corresponding to the simulation. (d) the difference between the observations and the simulation in panel (b). The overall shift upward in the simulation is due to increasing mean temperature. Most of the long timescale fluctuations in the bottom panel are due to changes in the seasonal cycle: at this location, temperatures are projected to warm more in winter months than in summer months. The short timescale fluctuations on the order of 0.1 degrees in the difference are due to changes in variability, which decreases in the future simulation. (e) marginal densities by season (labeled by corresponding months) for the observations in 2008-2009, the simulation from 2098-2099, the GCM from the same years under the historical forcing and the high future scenario scenario.

an argument for our procedure, which preserves features of the observations.

2.5.2 Assessing model fit and quality of emulation

To assess different aspects of how well our model describes the changes in covariance structure in CCSM3 in evolving climates, we show three diagnostics. First, we address in which geographic locations the model performs relatively better or worse. Second, we ask how well the statistical model performs as an emulator for a scenario on which the model has not been

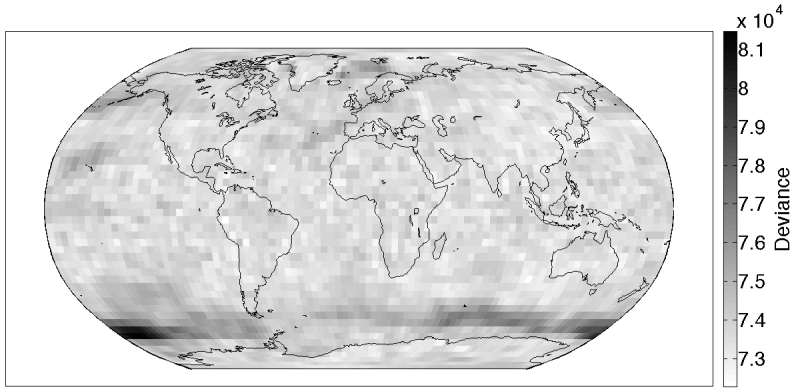

Figure 2.5: Deviances, comparing the approximate likelihood under our estimated model to that under the saturated model where the spectrum in each time block, scenario, and frequency has its own parameter. The number of observations at each location is about 1.6 million days. The regions showing the largest deviances are those where changes in sea ice have a strong influence on variability; in such locations, our model based on changes in mean temperature cannot be expected to be a fully adequate description of changes in variability.

trained. Finally, we show the extent to which the rate of change of mean warming improves the quality of emulation over the simpler model where changes in covariance are explained solely by changes in mean temperature.

To examine in which locations the model performs best and worst, we compare the deviances of our model at each location (Figure 2.5); recall that the deviance compares the likelihood under our estimated model to that under the saturated model where the value of the spectrum in each time block, scenario, and frequency is assigned its own parameter. The deviances are largest at the edge of the maximum present-day sea ice extent in the Southern Ocean, and relatively homogeneous elsewhere. The relatively poorer fit at ice margins is expected, since variability decreases substantially here as sea ice retreats, and those changes are therefore based in part on absolute temperatures. Any statistical model based purely on changes in temperature, rather than absolute temperatures, will have difficulties capturing variability changes due to phase changes between ice and water. This result should serve as a warning against using such methods over locations, scenarios, and time periods in which the response to $[CO_2]$ changes is highly nonlinear.

To address how well our statistical model is able to emulate GCM projected changes in variability for scenarios in some sense within the range spanned by our ensemble, we re-estimate our model using (a) all but the realizations under the medium scenario, and (b) only the realizations under the medium scenario. If the conclusions we draw from (a) match those drawn from (b), this is evidence that we have successfully emulated the changes in covariance structure under the medium scenario. We compare projected changes in marginal evolutionary spectra at year 2100 of the medium scenario, estimated under these two schemes (Figure 2.6). The estimated global patterns of changes in variability are quite similar under the two schemes. In an absolute sense, the biggest differences between the two schemes are at the lowest frequencies, but recall that since we have reason to believe that the ratios of spectra are less smooth at lower frequencies, we smooth with a smaller bandwidth at those frequencies and therefore our estimates of these changes are more uncertain. Globally, the differences between the estimates under schemes (a) and (b) are within two standard errors of zero in about 60-75% of the grid cells, depending on the frequency of interest. (Over land, the differences are within two standard errors in about 70-80% of the grid cells.) The locations where there are significant differences between the two schemes are, unsurprisingly, often those where we have argued that the model should have trouble, such as in the Southern Ocean. In these locations, the emulator usually underestimates variability changes.

As discussed above, a feature of our model is that changes in variability depend not only on the change in regional mean temperature but also on the rate at which those changes occur. One might ask whether the simpler model that omits the second term (i.e., $\delta_{l1}(\omega) = 0$) is just as good at emulating changes in variability. Figure 2.7 displays the predictive log likelihood ratio comparing the simpler model to our proposed model; by predictive log likelihood, we mean that the models are estimated as emulators, excluding the medium scenario realizations, and the likelihoods are evaluated for the medium scenario realizations (as such, no adjustment for model complexity is necessary). At all but three out of the 4,608 locations, the full model has a larger predictive likelihood than the simpler model, which

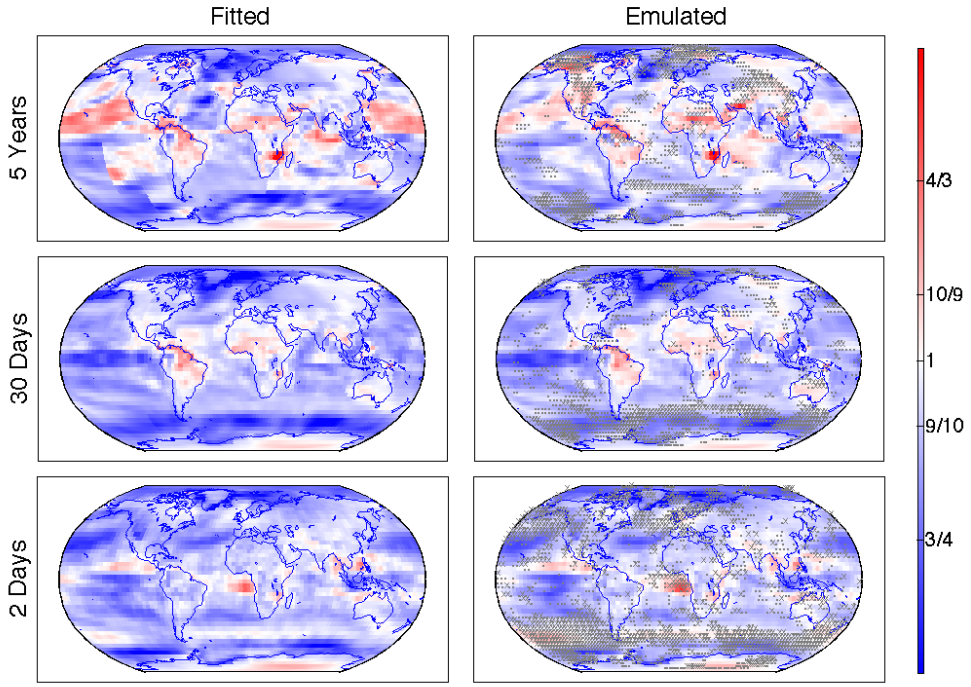

Figure 2.6: Estimates of changes in marginal spectra, at three frequencies, for the medium scenario at model year 2100 compared to the preindustrial climate. Left, estimates use only the medium scenario realizations. Right, estimates use all but the medium scenario realizations (i.e., are estimated as an emulator). Note that the apparent sudden change in behavior for low frequency variability in the Pacific Ocean (top left) is an artifact of the minimum bandwidth chosen for smoothing in these two adjacent regions; see Appendix 2.7.2 for details on bandwidth selection. Locations are marked with "." (or "x") when the difference between the emulator and the fitted model is more than two (or three) standard errors away from zero. The patterns are similar under both schemes, with most of the differences at locations where our model would not be expected to be a good description of changes in variability (e.g., at ice margins).

indicates that for the purposes of emulation it is useful to allow for a nonzero term involving the rate of change change of warming.

2.6 Discussion

In this work, we describe a method for transforming observed temperatures to produce simulations of future temperatures when the climate is in a transient state, based on the projected changes in means and temporal covariances in GCM output. We believe this

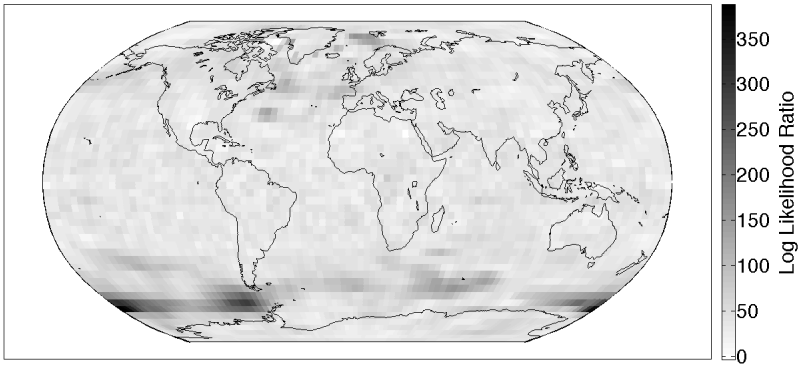

Figure 2.7: Log likelihood ratios comparing our proposed model to the reduced model taking $\delta_{l1}(\omega) = 0$. The models were estimated using all but the medium scenario realizations and the likelihoods were computed with just the medium scenario realizations, so the likelihood ratio is a comparison of the quality of emulation (no adjustment for model complexity is needed). In all but three locations on the globe, the likelihood under the full model is larger, which is an indication that for the purposes of emulation it is useful to include $\delta_{l1}(\omega)$ in the model.

approach should yield more realistic simulations of future climate than do either GCM runs or simulations based on modifying GCM runs.

Any observation-driven procedure is of course limited by the observational record. Observations have some intrinsic uncertainty not explicitly accounted for in this work. As suggested in Section 2.1, it would be possible to explore the effect of these uncertainties on the resulting simulation by repeating the simulation using different data products. Furthermore, the observational record is of a limited length and our procedure as described provides exactly one simulation of future temperatures equal in length to the observational record. As also suggested by Leeds et al. (2015), longer simulations could be produced either by recycling the observations entirely or by resampling them to generate new pseudo-observations.

An important feature of the procedure we describe is that our simulations preserve many features of the observational record not accounted for explicitly in the procedure. This chapter is concerned with changes in only the temporal covariance structure of temperatures in transient climates. In the methodology described here, the simulation therefore preserves,

for example, the spatial coherence spectra of the observations. While projected changes in spatial coherences in the model we study appear to be small (Section 2.7.5), changes in spatial coherences may also be important for societal impacts. We leave for future research the possibility of extending the methodology to account for such changes. This extension would be challenging and interesting, in part because temperatures are nonstationary in space with abrupt, local changes due to geographic effects.

Another challenging and interesting extension of our methods would be to jointly simulate future temperature and precipitation. Our work has focused on simulating temperatures, but potential changes in precipitation are also important for societal impacts. While there have been some model-driven proposals for jointly simulating temperature and precipitation (Piani and Haerter, 2012; Vrac and Friederichs, 2015), to our knowledge most approaches (both model- and observation-driven) proceed by simulating the two quantities separately. Versions of the simple Delta method can and are used for monthly precipitation, with the assumption that the GCM captures multiplicative (rather than additive) changes in rainfall amount (see, e.g., Teutschbein and Seibert (2012) and references therein for a review of common precipitation simulation methods). A simple Delta method cannot, however, capture the changes on the timescale of individual rainfall events, whose intensity changes differently than that of time-averaged rainfall, with projections of less frequent but more intense storms (e.g., Trenberth (2011)). Our approach, based on spectral methods, is likely also inadequate for characterizing changes in variability of daily precipitation, because daily precipitation often takes the value zero. (Data sources for daily local precipitation are also more problematic than those for temperatures, since reanalyses provide a poor reproduction of precipitation structure and interpolating sparse, complex precipitation fields is difficult.) More sophisticated observation-driven simulation methods for precipitation remain an area of research, as does the joint simulation of temperature and precipitation. In the context of this chapter, where changes to the correlation structure of temperatures are detectable but not very large, we expect that separately simulating temperature and precipitation (as in

common practice) will not result in large changes to their bivariate dependence structure.

Finally, the characterization of variability changes in transient climates is itself an issue of scientific interest. One of our motivations for developing the methods described here is to enable studying and ultimately comparing different GCM projections of changes in temperature variability. Most publicly available GCM runs (including most runs mandated by the IPCC) describe plausible future climates, which are necessarily in transient states. The statistical model we develop uses the GCM change in regional mean temperature and its rate of change to describe the GCM projected changes in covariance structure; in the GCM runs we study, these factors effectively summarize the projected changes in covariance. While we have investigated changes in variability in only one GCM, at relatively coarse resolution, we hope that our methods are applicable across GCMs, and may aid in carrying out a comparison across different GCMs in a coherent and interpretable way.

2.7 Appendix

2.7.1 Estimating changes in regional and local mean temperature

The model for changes in covariance, (2.6), requires an estimate of the changes in regional mean temperature, $\bar{\Delta}_S^{(s,B)}(t)$, for scenario s. We also need changes in local mean temperature to compute the simulation (2.4). We estimate these using a modification of the mean emulator described in Castruccio et al. (2014). Write $\bar{\mu}_S^{(s)}(t)$ for the regional mean temperature at time t under scenario s in region S. We assume that

$$\begin{aligned}\bar{\mu}_S^{(s)}(t) = \quad & \beta_{0,S} + \beta_{1,S} C^{(s)}(t) + \sum_{k=1}^{K} \left\{ \gamma_{k,S} \cos\left(\frac{2\pi t k}{365}\right) + \zeta_{k,S} \sin\left(\frac{2\pi t k}{365}\right) \right\} \\ & + \sum_{k=1}^{K} \left\{ \gamma'_{k,S} C^{(s)}(t) \cos\left(\frac{2\pi t k}{365}\right) + \zeta'_{k,S} C^{(s)}(t) \sin\left(\frac{2\pi t k}{365}\right) \right\}\end{aligned} \tag{2.14}$$

where

$$C^{(s)}(t) = \sum_{m=0}^{\infty} (1-\phi_S)\phi_S^m \log\left(\frac{[\mathrm{CO_2}]^{(s)}(t-m)}{[\mathrm{CO_2}]^{(\mathrm{B})}}\right)$$

for $\phi_S \in [0,1)$, and $[\mathrm{CO_2}]^{(s)}(t)$ and $[\mathrm{CO_2}]^{(\mathrm{B})}$ are the $\mathrm{CO_2}$ concentrations under scenario s at time t and under preindustrial conditions, respectively. For the harmonic terms, we take $K = 6$. This mean emulator differs from that described in Castruccio et al. (2014) in essentially two ways. First, here we exclude one effect in theirs that was meant to distinguish short term and long term effects of changes in $[\mathrm{CO_2}]$. For smooth, monotonic scenarios like those in our ensemble, it is difficult to distinguish these two effects. Second, whereas their paper was concerned with emulating annual average temperatures, here we are interested in daily temperatures, so we need terms that account for the (possibly changing) mean seasonal cycle; note that the GCM runs use a year of exactly 365 days, hence our representation of the seasonal cycle in terms of harmonics of 1/365. In most regions, any changes to the mean seasonal cycle are small besides an overall increase in mean, although in regions with strong seasonal cycles, the mean seasonal cycle tends to be damped more in winter months than summer months (Figure 2.8).

For the change in regional mean temperature used as an input in (2.6), we define $\bar{\Delta}_S^{(s,B)}(t) = \beta_{1,S} C^{(s)}(t)$, which is, under the above model, the change in regional mean temperature from the preindustrial climate excluding changes in the seasonal cycle. For the changes in local mean temperature used as an input in the simulation (2.4), we assume that the local means are related to the regional means through a regional pattern scaling relationship; that is, the change in local mean from the preindustrial climate is taken to be proportional to the change in regional mean from preindustrial climate:

$$\tilde{\mu}_l^{(s)}(t) - \tilde{\mu}_l^{(B)} = \lambda_l(\bar{\mu}_S^{(s)}(t) - \bar{\mu}_S^{(B)}).$$

This is the approach advocated by Castruccio et al. (2014) for grid cell-level mean emulation.

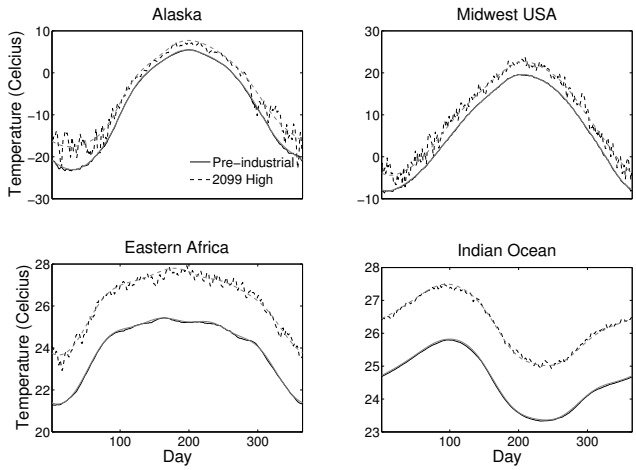

Figure 2.8: Examples of emulated (gray) and empirical (black) mean seasonal cycles in four regions compared between the preindustrial climate (solid) and the climate under the high scenario in the year 2099 (dashed). The emulations are fitted according to (2.14). The empirical estimates are obtained by averaging across grid cells in the region of interest and across time (for the equilibrated preindustrial run) or across realizations (for the high scenario realizations). The empirical estimates of the 2099 seasonal cycles (red) are noisy because they represent averages across only eight realizations of the regional mean temperature.

2.7.2 Choice of variable bandwidth for estimating $\delta_{li}(\omega)$

Here we describe the choice of weights, $w_{k,j,i}$, used in the estimates of $\delta_{l,i}^{(s)}(\omega_j)$ given by (2.13). We let $w_{k,j,i}$ correspond to a variable-bandwidth quadratic kernel,

$$w_{k,j,i} \propto \begin{cases} 1 - (\frac{k}{M_{j,i}+1})^2, & k \in \{-M_{j,i}, ..., M_{j,i}\} \\ 0, & \text{otherwise} \end{cases}$$

where $M_{j,i}$ controls the bandwidth at frequency ω_j for $i = 0, 1$, which we allow to vary as $M_{j,i} = (m_{1,i} - m_{0,i})h_j(p_i) + m_{0,i}$ where $m_{1,i}$ and $m_{0,i}$ are such that $0 < m_{0,i} \leq m_{1,i}$ and

$$h_j(p_i) = \begin{cases} \frac{1}{2}(1 + \cos\left(\frac{2\pi}{p_i}(\omega_j - \frac{p_i}{2})\right)), & 0 \leq \omega_j < \frac{p_i}{2} \\ 1 & p_i \leq \omega_j \leq 1 - p_i \\ \frac{1}{2}(1 + \cos\left(\frac{2\pi}{p_i}(\omega_j - \frac{p_i}{2})\right)) & 1 - \frac{p_i}{2} < \omega_j < 1, \end{cases}$$

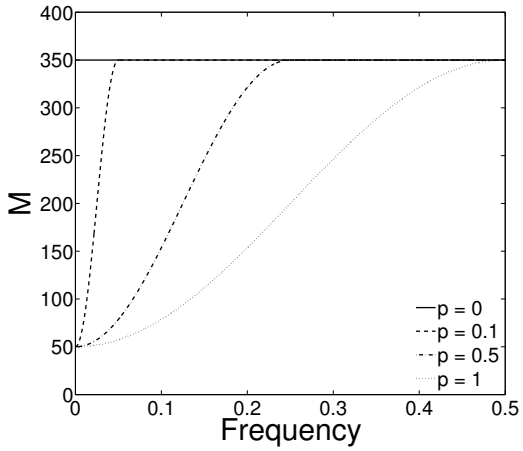

Figure 2.9: Examples of the function $M_{j,i}$ controlling the bandwidth of the kernel estimator as a function of the frequency being estimated. In each of these curves $m_{0,i} = 50$ (controlling the minimum bandwidth) and $m_{1,i} = 350$ (controlling the maximum bandwidth), and the value of p_i (controlling the transition between the minimum and maximum bandwidths) is noted in the figure. The form of $M_{j,i}$ is such that the bandwidth of the kernel is smaller at low frequencies.

so $m_{1,i}$ controls the maximum bandwidth, $m_{0,i}$ the minimum bandwidth, and p_i the frequency at which the bandwidths transition between $m_{0,i}$ and $m_{1,i}$ (see Figure 2.7.2). The reason we allow the bandwidths to decrease in frequency in this constrained way is that we have found that the log ratio of spectra in our ensemble are typically less smooth at very low frequencies. The specific form of h is somewhat arbitrary, but appears to work well in practice.

To implement the estimator, we need to choose the parameters $m_{1,i}$, $m_{0,i}$, and p_i. This is done via leave-one-out cross validation; that is, for each region the parameters are chosen to minimize the sum of the cross-validation scores for each location,

$$R_{CV}(m_{1,i}, m_{0,i}, p_i) = \sum_{j,l} \frac{(\hat{\delta}_{li}(\omega_j) - \delta^*_{li}(\omega_j))^2}{(1 - w_{0,j,i})^2},$$

where δ^*_{li} is the maximizer of the approximate likelihood, (2.11), and $\hat{\delta}_{li}$ is the smoothed version with weights $w_{k,j,l}$ defined above. These parameters vary by region, but typical

values of $(m_{1,i}, m_{0,i}, p_i)$ when estimating δ_{l0} and δ_{l1} using all runs in our ensemble are around $(350, 50, 0.4)$ for $i = 0$ and $(800, 400, 0.7)$ for $i = 1$ (recall that the local periodograms in (2.11) are taken over ten-year blocks, so the raw estimates of δ_{li} are calculated at 1,825 frequencies). Since the parameters vary regionally, there may be sudden changes in estimates of the changes in spectra near boundaries of the region (see Figure 2.6, top right). One could consider a post-hoc spatial smoothing of either the bandwidths or the changes in spectra to avoid this, but this is not explored in this chapter.

2.7.3 Standard errors for estimates of $\delta_{li}(\omega)$

Under (2.11), the Fisher information for $(\delta_{l0}, \delta_{l1}, \tilde{a}^{(B)})$ is a block diagonal matrix with blocks for each frequency, ω_j, approximately (under standard approximations for periodograms) equal to

$$\mathcal{I}(\omega_j) = \begin{bmatrix} \sum_{sb}(\bar{\Delta}_b^{(s,B)})^2 R'_s & \sum_{sb}(\partial_t \bar{\Delta}_b^{(s,B)})\bar{\Delta}_b^{(s,B)} R'_s & \sum_{sb} \frac{\bar{\Delta}_b^{(s,B)} R'_s}{\tilde{a}^{(B)}(\omega_j)} \\ & \sum_{sb}(\partial_t \bar{\Delta}_b^{(s,B)})^2 R'_s & \sum_{sb} \frac{\partial_t \bar{\Delta}_b^{(s,B)} R'_s}{\tilde{a^{(B)}}(\omega_j)} \\ & & \sum_{sb} \frac{R'_s + M}{\tilde{a}^{(B)}(\omega_j)^2}, \end{bmatrix} \tag{2.15}$$

where $R'_s = R_s - 1$. We calculate standard errors under the usual approximation that the variance of the maximum likelihood estimate is the inverse information matrix; such an approximation can be justified by, for example, considering either the number of independent GCM runs or the number of time blocks to be large. The resulting covariance matrix for $\delta^*_{l0}(\omega_j)$ and $\delta^*_{l1}(\omega_j)$, the maximum likelihood estimators for the δ functions, is constant in

frequency. Write

$$\begin{aligned}\operatorname{Var}\begin{pmatrix}\delta^*_{l0}(\omega_j)\\ \delta^*_{l1}(\omega_j)\end{pmatrix} &\approx \begin{bmatrix}\mathcal{I}(\omega_j)^{-1}_{1,1} & \mathcal{I}(\omega_j)^{-1}_{1,2}\\ & \mathcal{I}(\omega_j)^{-1}_{2,2}\end{bmatrix}\\ &\equiv \begin{bmatrix}V_{\delta_{l0},\delta_{l0}} & V_{\delta_{l0},\delta_{l1}}\\ & V_{\delta_{l1},\delta_{l1}}\end{bmatrix}.\end{aligned}$$

Then the variance of the smoothed estimates, $\hat{\delta}$, defined in (2.13), at frequencies not close to 0 or π is

$$\begin{aligned}\operatorname{Var}\begin{pmatrix}\hat{\delta}_{l0}(\omega_j)\\ \hat{\delta}_{l1}(\omega_j)\end{pmatrix} &\equiv \begin{bmatrix}V_{\hat{\delta}_{l0},\hat{\delta}_{l0}} & V_{\hat{\delta}_{l0},\hat{\delta}_{l1}}\\ & V_{\hat{\delta}_{l1},\hat{\delta}_{l1}}\end{bmatrix}\\ &= \begin{bmatrix}V_{\delta_{l0},\delta_{l0}}\sum_k w^2_{k,j,0} & V_{\delta_{l0},\delta_{l1}}\sum_k w_{k,j,0}w_{k,j,1}\\ & V_{\delta_{l1},\delta_{l1}}\sum_k w^2_{k,j,1}\end{bmatrix}.\end{aligned}$$

When 0 or π is within the local bandwidth for the frequency of interest, we make the correction to the above to account for the fact that $\delta^*_{li}(\omega_j)$ is periodic and even (here omitted for simplicity). The variance of $\log\hat{\rho}^{(s,B)}_l(t,\omega)$, the smoothed estimate of log ratio of spectra, is then approximately

$$\begin{aligned}\operatorname{Var}\log\hat{\rho}^{(s,B)}_l(t,\omega) \approx (\bar{\Delta}^{(s,B)}_S(t))^2 V_{\hat{\delta}_{l0},\hat{\delta}_{l0}} + (\partial_t\bar{\Delta}^{(s,B)}_S(t))^2 V_{\hat{\delta}_{l1},\hat{\delta}_{l1}}\\ +2\bar{\Delta}^{(s,B)}_S(t)\partial_t\bar{\Delta}^{(s,B)}_S(t)V_{\hat{\delta}_{l0},\hat{\delta}_{l1}}.\end{aligned}$$

2.7.4 Computing simulations

Computing (2.4) efficiently requires the ability to quickly compute the products

$$C_{N_T}\left(\sqrt{\hat{\rho}^{(s,B)}}\right)x \text{ and } C^{-1}_{N_T}\left(\sqrt{\hat{\rho}^{(0,B)}}\right)x$$

for a vector x. The matrix-vector products may be written as

$$\begin{aligned}\left(C_{N_T}\left(\sqrt{\hat{\rho}^{(s,B)}}\right)x\right)_t &= \sqrt{\frac{2\pi}{N_T}}\sum_{j=0}^{N_T-1} x_j e^{\frac{1}{2}[\bar{\Delta}_t^{(s,B)}\hat{\delta}_{l0}(\omega_j)+\partial_t\bar{\Delta}_t^{(s,B)}\hat{\delta}l1(\omega_j)]}e^{i\omega_j t} \\ &= \sqrt{\frac{2\pi}{N_T}}\sum_{j=0}^{N_T-1} x_j \sum_{p=0}^{\infty}\frac{\{\frac{1}{2}[\bar{\Delta}_t^{(s,B)}\hat{\delta}_{l0}(\omega_j)+\partial_t\bar{\Delta}_t^{(s,B)}\hat{\delta}_{l1}(\omega_j)]\}^p}{p!}e^{i\omega_j t} \\ \approx \sqrt{\frac{2\pi}{N_T}}\sum_{p=0}^{P-1}\sum_{m=0}^{p}&\frac{(\bar{\Delta}_t^{(s,B)})^{p-m}(\partial_t\bar{\Delta}_t^{(s,B)})^m}{2^p m!(p-m)!}\sum_{j=0}^{N_T-1}\hat{\delta}_{l0}(\omega_j)^{p-m}\hat{\delta}_{l1}(\omega_j)^m x_j e^{i\omega_j t}\end{aligned}$$

where the approximation just truncates the Taylor series at $P-1$, then uses the binomial expansion and changes the order of summation. This approximation is the weighted sum of $P(P+1)/2$ inverse discrete Fourier transforms; if P can be taken to be only modestly large so that the Taylor approximation is accurate, this sum can be computed efficiently. In our setting, the changes in variability projected by the GCM are small and P need not be very large in order for the approximation to be quite good; we have found that taking $P = 10$ is more than enough to give accurate approximations given the magnitude of the estimated changes in variability.

To compute the matrix-inverse vector products, we have found, as in Guinness and Stein (2013b), that iterative methods work well. In order to work well, these require the ability to compute forward multiplication quickly, which we have just described, as well as a good preconditioner; since the historical time series is only mildly nonstationary, we have found that the Fourier transform scaled by the square root of the average of $\rho^{(0,B)}(t,\omega)$ over time works well as a preconditioner.

2.7.5 Additional supplementary figures

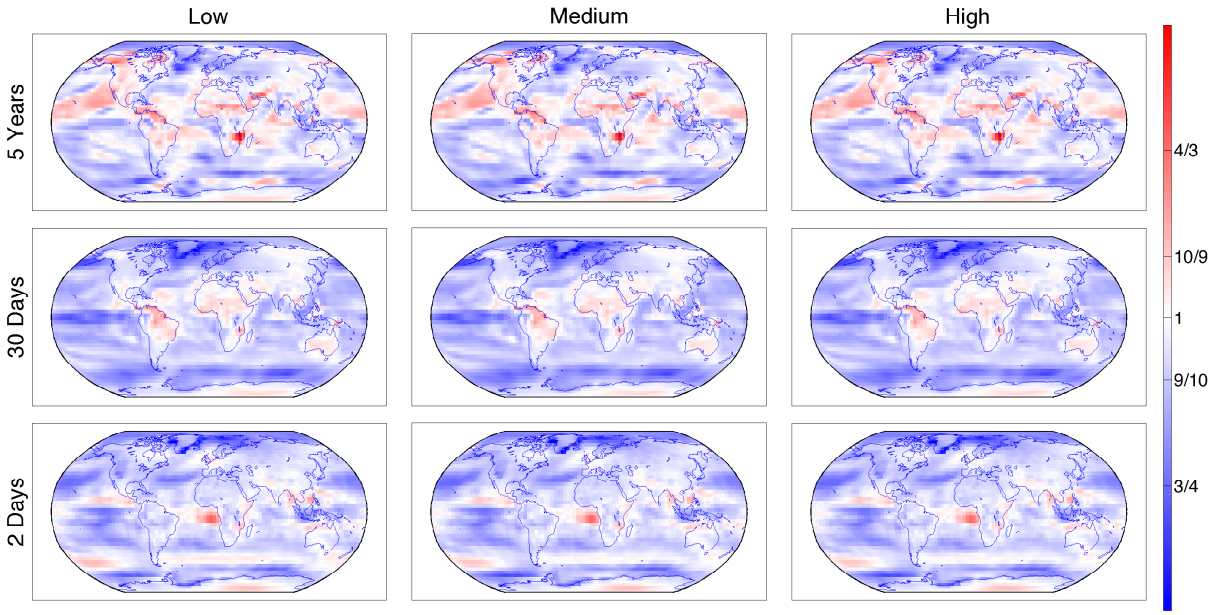

Figure 2.10: Estimates of changes in marginal spectra, at three frequencies, at the year corresponding to that with the same regional mean temperature as in the low scenario in model year 2100 (i.e., left, $\rho_l^{(L,B)}(t,\omega)$ at year 2100; middle, $\rho_l^{(M,B)}(t,\omega)$ at years ranging from 2050-2056 depending on region; right, $\rho_l^{(H,B)}(t,\omega)$ at years ranging from 2037-2044 depending on region). Red indicates an increase in variability and blue a decrease in variability. The estimated differences between scenarios are small because the effect of the rate of change of warming on the estimated changes in variability are smaller than the effect of the overall regional mean change in temperature. A direct comparison of the estimates between the high and low scenario is shown in Figure 2.3 of the main text.

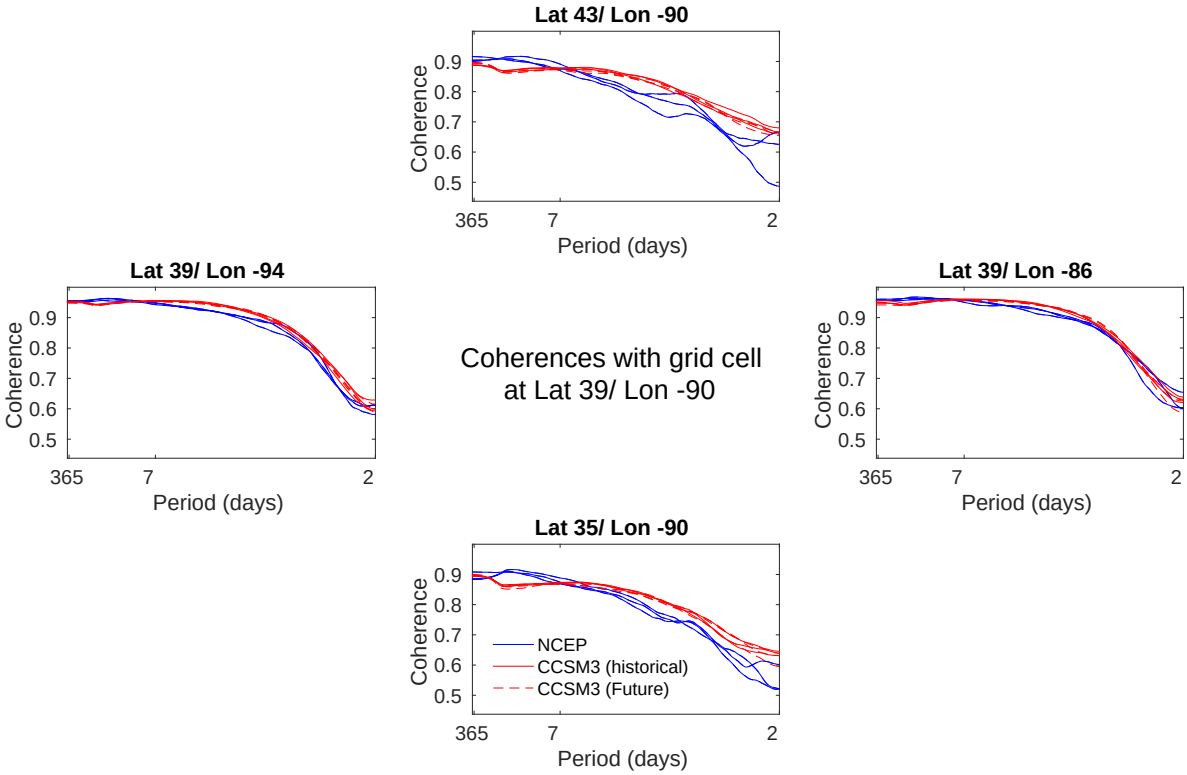

Figure 2.11: Estimates of coherence spectra between temperatures in the Midwestern grid cell discussed in Section 2.5.1 and its four neighboring grid cells; the coordinates corresponding to each neighbor are indicated (and figures are positioned geographically). Estimates are shown for the ten-year blocks 1980-1989, 1990-1999, and 2000-2009 as well as the corresponding time 90 years in the future under the high scenario (i.e., 2070-2079, 2080-2089, and 2090-2099). Estimates in red correspond to temperatures from CCSM3 under historical forcing (solid) and in the high scenario future (dashed). Those in blue correspond to temperatures from NCEP CFSR and the simulation (the coherences do not change in the simulation). The time series were processed to remove means and (marginal) seasonal variability as described in the main text. Marginal spectra, co-, and quadrature spectra were estimated by averaging the raw estimates for each run (for the CCSM3 runs, where we have multiple runs) and then smoothing each component using a quadratic kernel; since we cannot average over multiple runs of reanalysis, the estimates for reanalysis were smoothed using a slightly wider bandwidth. While the coherence spectra are arguably similar in the East-West direction, CCSM3 has temperatures that are more coherent than the reanalysis's in the North-South direction. Coherences do not change much in the future GCM run, which suggests that our approach for the simulations may be adequate.

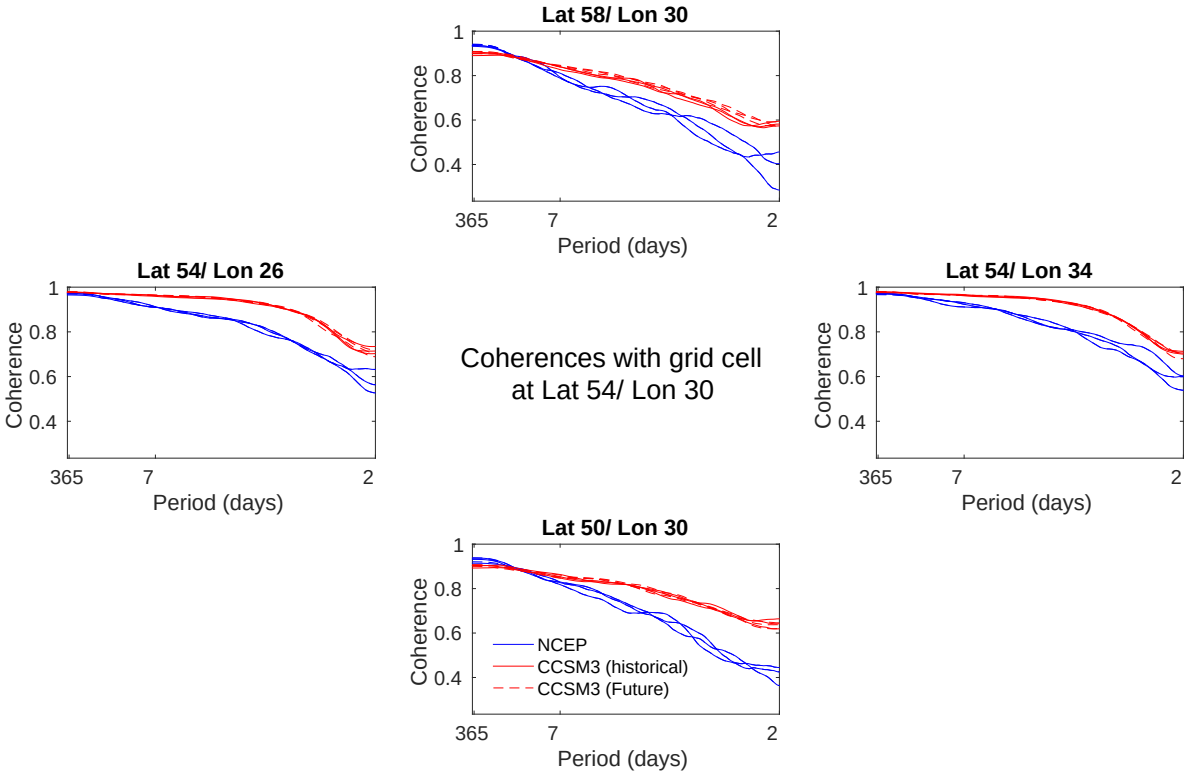

Figure 2.12: The same as Figure 2.11 but for a central grid cell in Northern Europe. Temperatures are more coherent in CCSM3 than in the reanalysis.

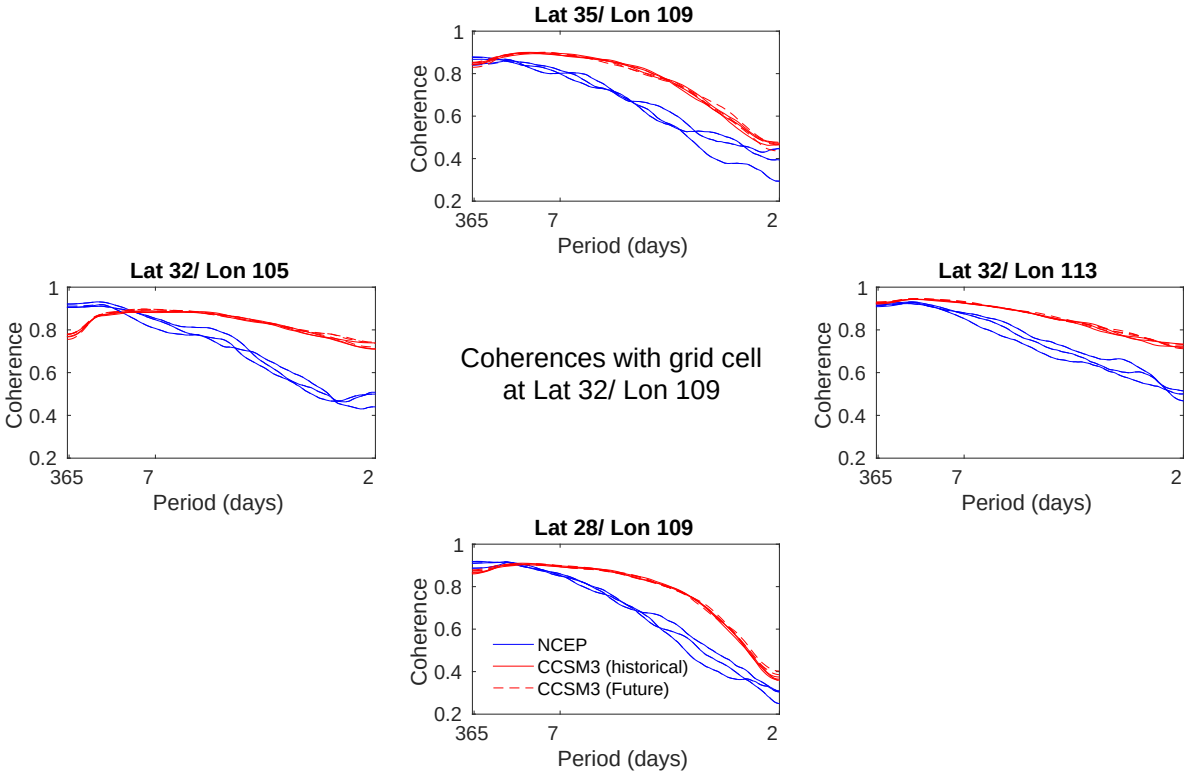

Figure 2.13: The same as Figure 2.11 but for a central grid cell in East Asia. Temperatures are more coherent in CCSM3 than in the reanalysis.

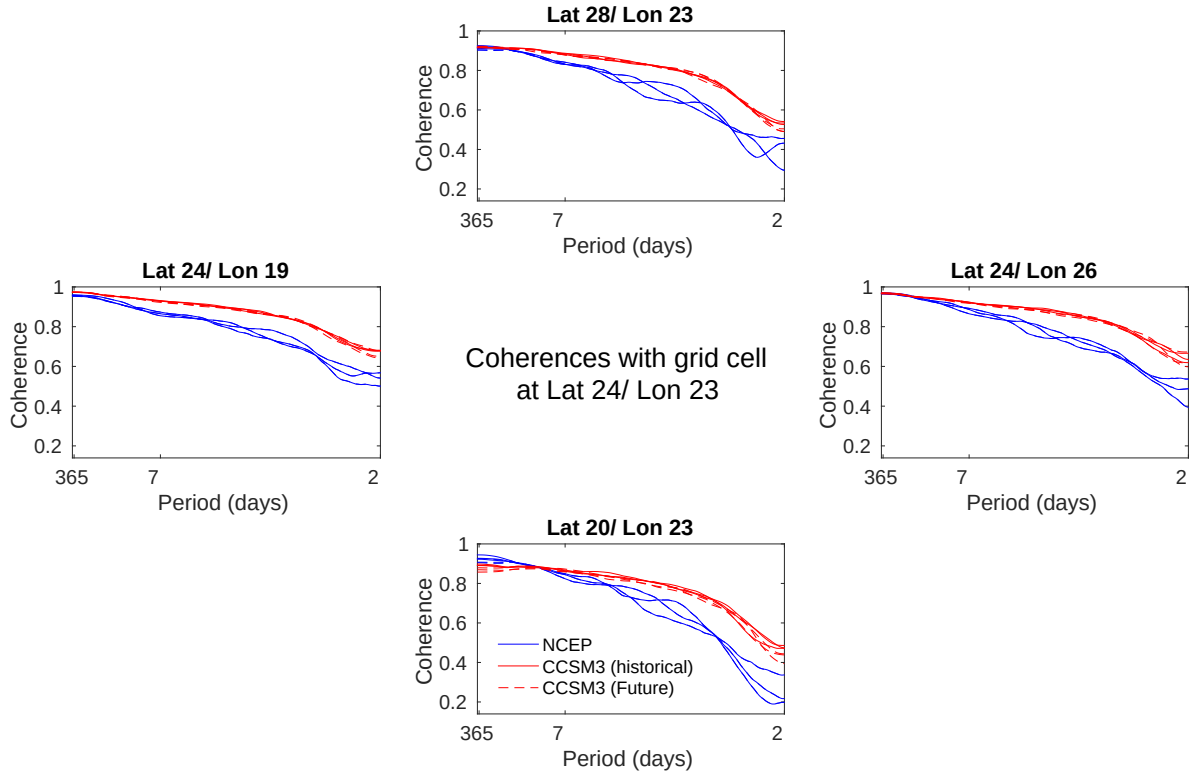

Figure 2.14: The same as Figure 2.11 but for a central grid cell in the Sahara. Temperatures are more coherent in CCSM3 than in reanalysis.

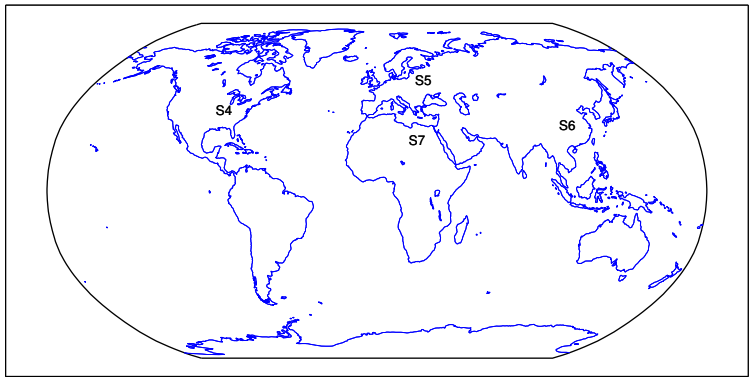

Figure 2.15: The locations of the central grid cells corresponding to the coherency plots in Figures 2.11 - 2.14 (labeled by figure number). The grid cell for Figure 2.11 corresponds to the one discussed in Section 2.5.1 of the main text.

CHAPTER 3
USING COVARIATES TO MODEL DEPENDENCE IN NONSTATIONARY, HIGH-FREQUENCY METEOROLOGICAL PROCESSES

3.1 Introduction

Over the past two decades, a wealth of data has become available providing measurements of surface meteorological quantities at very fine timescales. This data has offered new opportunities to learn about the high-frequency properties of these processes – both marginally, and potentially as a system of many dependent processes. While such data can be a rich source of information, meteorological processes often exhibit behaviors that present the statistical modeler with serious challenges. We here discuss a small subset of these challenges, focusing on the temporal behavior of a nonstationary bivariate process – the high frequency changes in surface temperature and dew point – and how that behavior changes with covariates.

The joint behavior of temperature and dew point presents an interesting statistical challenge because we may expect the joint behavior to be different when temperature is close to the dew point compared to when not, since dew point is bounded above by temperature. (Stated equivalently, we may expect the behavior to be different when relative humidity is close to 100% than when not.) It has also been previously noted that the high-frequency variability of temperature appears to vary with, for example, solar radiation (Guinness and Stein, 2013a). These observations suggest that covariate information can help explain the high frequency behavior of the process of interest. In this chapter, we develop a model for the changes in the high-frequency behavior of temperature and dew point conditionally on the level of relative humidity and on other measured quantities.

Methods for modeling a single meteorological quantity at high temporal frequencies have been explored by Stein (2009), Chang and Stein (2014), and Guinness and Stein (2013a),

where the goal has been spatial interpolation and the quantity modeled has been either surface temperature or pressure. Since these processes clearly exhibit temporal nonstationarity, a part of this previous work has been to cope with this nonstationarity, either by estimating changes in the variability of pressure in time using spatial information (Stein, 2009; Chang and Stein, 2014) or by modeling the evolutionary spectrum of the temperature process (Guinness and Stein, 2013a); for information on evolutionary spectra, see Priestley (1981). In contrast, our focus is on modeling a bivariate process in time and we seek to explain its nonstationary behavior with covariates.

A number of approaches have been proposed for bivariate geostatistical models; see, for example, Fanshawe and Diggle (2012) for a review. Since constructing valid cross-covariance functions is not straightforward, most of the dominant approaches have represented the bivariate process as a linear combination of independent processes (Goulard and Voltz, 1992) or as a convolution of a process with a set of kernel functions (Ver Hoef and Barry, 1998), although Gneiting et al. (2010) is an example of a parametric model for the cross-covariance function being directly provided. There are other well-known methods for modeling multivariate linear time series in both the time and frequency domains; for background, see Shumway and Stoffer (2011). A rich treatment of spectral methods in particular is given in Priestley (1981).

Still other methods have been put forth for incorporating covariates into nonstationary covariance models (Calder, 2008; Schmidt et al., 2011; Neto et al., 2014; Reich et al., 2011; Ingebrigtsen et al., 2014). A literature review is given in Neto et al. (2014). The approaches have been to begin with a nonstationary model and to find a way to incorporate covariates into whatever drives the nonstationarity – for example, a latent space model with covariate information in the latent space (Schmidt et al., 2011), or a linear combination of stationary processes where the weights on each process are a function of covariates (Reich et al., 2011). Methods for covariate dependent spectral analysis have also been explored in the time series literature; see, for example, Krafty et al. (2012) and references therein. Here it is often

assumed that observations are multiple independent realizations of time series with spectra that are functions of covariates. A fully nonparametric approach for univariate (categorical or quantitative) time series based on local smoothing is discussed in Krafty et al. (2012); here we favor a parametric approach, in some measure because of its interpretability.

We propose a model specifying a locally stationary process that has a bivariate spectral matrix that varies parametrically with covariates over blocks of time, providing a description of how the spectral matrix changes as covariates change. The models and methods we propose have been chosen with the particular features of the process of interest in mind. Since many of these features are arguably peculiar to high-frequency meteorological processes, we favor an approach that has this specific process in mind, as opposed to more general concerns. That stated, we view frequency-domain models as being a natural choice for bivariate processes in time; in addition to being easily interpretable, such models sidestep some of the difficulties in directly building valid bivariate covariance models.

The remainder of this chapter discusses the temperature and dew point data and the model we consider. In Section 3.2, we give an exploratory analysis, with a discussion of how the behavior of temperature and dew point appears to change with covariates. In Section 3.3, we describe our model, and in Section 3.4 we discuss the results from model fitting, including a criterion for model comparison and model diagnostics when spectra depend on covariates. Section 3.5 provides some concluding remarks. Supplementary material is given in Section 3.6.

3.2 Exploratory Analysis

The data analyzed are from the Atmospheric Radiation Measurement (ARM) Program's Southern Great Plains (SGP) site in Northern Oklahoma, at the central facility near Lamont, OK. Measurements are taken at one-minute intervals, with temperature reported to the nearest hundredth of a degree Celsius and relative humidity to the nearest tenth (or sometimes hundredth) of a percentage point. We focus our discussion on measurements taken

over May 1-5 from the years 2003 through 2012. Two days in this period – May 1, 2003 and May 5, 2006 – were removed prior to analysis due to missing values, for a total of 48 days (i.e. 69,120 minutes) included in the analysis. Removing these two full days rather than, for example, only the hours with missing values has little effect on the analysis described in the following section, but somewhat simplifies some of the procedures we describe. We chose the relatively short period of time in the beginning of May to avoid challenges related to seasonality. While we have not examined the data at other times of year, this process will clearly exhibit substantial differences across seasons.

We consider the temperature and dew point processes. Dew point itself is not measured directly, but can be approximated using the temperature and relative humidity measurements according to, for example, the approximation given in Eq. (8) of Lawrence (2005). Many notable features can be seen in a raw plot of the data, the first day of which we present in the top left part of Figure 3.1. The main features to recognize are: (i) dew point is bounded above by temperature; (ii) temperature and dew point often nearly coincide, especially throughout the early morning; and (iii) the high-frequency variability of the processes varies throughout the day and is typically greater in the daylight portion of the day.

These features may be thought of as existing on different scales: the bounding of dew point by temperature may be characterized by the low frequency behavior of the temperature and dew point process, and the changes in variability by the behavior of the rough, residual process. To explore this, both the temperature and dew point were smoothed with a three-hour moving average and the residual component was computed as the difference between the observed data and the smoothed data. These components are displayed in the bottom left of Figure 3.1. While a three-hour moving average is employed here, we have explored the sensitivity of the analysis described in the following sections to this choice and have found that the length of the smoothing window does not have a large effect on the results. More detailed information about the effect of this choice is provided in the Appendix 3.6.1.

Note that if dew point is bounded above by temperature, this constrains the behavior

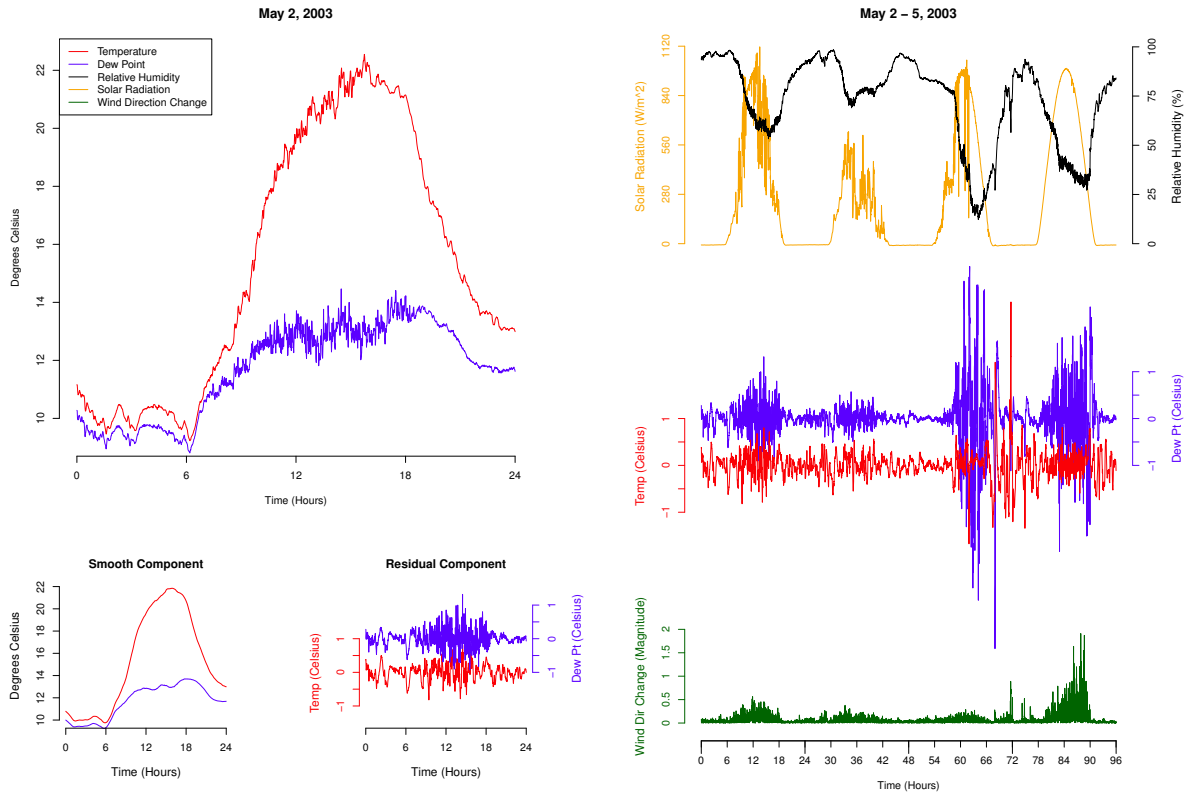

Figure 3.1: Left: Temperature and dew point on the first day included in the data analyzed, and decomposed into smoothed and residual components. Right: level of relative humidity and solar radiation, residual components of temperature and dew point, and magnitude of changes in wind direction. All three covariates appear to help explain the variation in variability of the residual processes, with variability tending to be higher in daylight hours, when relative humidity is low, and when the minute-to-minute changes in wind direction are large.

of the residual processes in addition to the smooth process; however, if we allow that temperature can fall modestly below dew point for a short period of time, then the bounding issue can essentially be relegated to the smooth component as long as the variability of the difference between the temperature and dew point residuals goes to zero as the smooth component of relative humidity approaches 100%. While the instrument measuring relative humidity enforces the constraint that relative humidity be below 100% – and therefore that dew point be below temperature – our model, described in the Section 3.3, will implicitly allow for small violations of this constraint.

Consider, then, the residual component. A salient feature of this component is that

its changes in variability can apparently be explained by other observed covariates. Both downwelling shortwave (i.e., solar) radiation and the relative humidity appear to be related to the level of the variability of the residual components, with the variability apparently increasing as solar radiation increases and as relative humidity decreases (Figure 3.1, top right). This relationship is apparent even beyond the mere fact that all three typically have a diurnal cycle. Since solar radiation and relative humidity are strongly correlated, it is difficult to determine which, if not both, of the two processes is driving the changes in high-frequency variability; relative humidity tends to be high when solar radiation is low. The level of relative humidity, however, seems to explain changes in variability even in nighttime hours when solar radiation is zero. The explanation postulated in Guinness and Stein (2013a) is that convective forces drive the relationship between the changes in variability and solar radiation. We additionally argue that when relative humidity is very high, its variability should be lower in order for the relative humidity to remain below 100% – and thus as temperature approaches dew point, the difference between temperature and dew point will be less volatile.

A third covariate that apparently explains some of the changes in the distributional properties of the residual component is the wind direction. We have found that the high-frequency volatility of the temperature and dew point residual processes is greater when the minute-to-minute changes in the wind direction are greater (Figure 3.1, bottom right). Our measure of change in wind direction is defined in Section 3.3. As with relative humidity and solar radiation, the relationship between the magnitude of wind direction changes and the volatility of the residual processes apparently extends beyond the typical diurnal cycle observed in both. Here, we argue that the mechanism driving the changes in volatility is the bulk movement of air mass; if the air is moving in many directions, its temperature and dew point at a fixed location will tend to be more volatile. Wind direction has also been used in Neto et al. (2014) to model the nonstationary spatial behavior of ground-level ozone and in Calder (2008) to model the nonstationary spatial behavior of particulate matter in the

ambient air.

The remainder of our analysis assumes that the changes in the features of the residual components can be explained solely by the relative humidity, wind direction, and presence of daylight. In reality, other processes will exert some influence on the changes in volatility of these processes. In particular, three other quantities are measured concurrently as part of the ARM SGP surface meteorology datastream: wind speed, pressure, and precipitation. In our exploratory analysis, the changes in volatility were apparently less related to these covariates and so were left out of the ensuing analysis. For example, since one might expect the behavior of temperature and dew point to be different in rainy and non-rainy conditions, the ensuing analysis was performed both on all hours and excluding hours with precipitation. The results from the latter were not found to be substantially different, so we provide results from fitting the model to all of the data.

3.3 Model

We provide a statistical model for the joint behavior of the residual temperature and dew point processes conditionally on the smooth component of the process and other covariates, which we will herein assume has already been observed or modeled. Because we want to enforce the constraint that temperature can only fall mildly below dew point, we find it more natural to model the temperature residual and the difference between the temperature and dew point residuals, which we will call the difference residual. The difference between temperature and dew point is also called the dew point spread and hence the difference residual may also be called the dew point spread residual.

Let $T_t^{(0)}$ be the temperature residual at time t for $t = 1, ..., N$; $D_t^{(0)}$ be the same for dew point; and $\Delta_t^{(0)} = T_t^{(0)} - D_t^{(0)}$ be the difference residual. Let $T_t = T_{t+1}^{(0)} - T_t^{(0)}$ be the first differences of $T_t^{(0)}$ and similarly define D_t and Δ_t. Finally, let X_h be a collection of covariates observed in a block of time of length L, $\mathbf{h} = \{L(h-1)+1, L(h-1)+2, ..., Lh\}$, indexed by the block number, $h \in \{1, ..., H\}$, where H is the total number of blocks. We

model $\{(\Delta_t, T_t)\}_{t=1}^{N}$ conditionally on $\{X_h\}_{h=1}^{H}$.

Our approach assumes that (i) the process is approximately stationary within each $\mathbf{h}$ for suitably small L, and (ii) the non-stationarity across the blocks is driven entirely by X_h. We view the stationarity assumption as reasonable if L is small enough. Ultimately, we will argue that the model successfully captures the changes in short timescale volatility observed in the data, and to the extent that this is true, assumption (ii) may be regarded as reasonable. Possible relaxations of these assumptions are discussed in Section 3.5.

Below, we take the length of the hourly blocks to be $L = 60$ minutes. We have also performed the analysis using both longer and shorter blocks within which the process is assumed stationary. We have found that varying the length of the blocks between 30 minutes and two hours does not lead to substantially different estimates. More information on this sensitivity analysis is provided in the Appendix 3.6.1.

3.3.1 Model for bivariate spectrum

Given our assumptions, it is natural to model the bivariate process as having, within each block of time, a bivariate spectral representation that is a function of X_h. Recall that a mean zero Gaussian process can be completely described by two nonnegative marginal spectra, a coherence between zero and one, and a phase spectrum; as such, specifying a model in the spectral domain rather than the time domain sidesteps some difficulties in constructing valid and interpretable bivariate covariance functions.

More specifically, suppose that (Δ_t, T_t) is a mean zero Gaussian process satisfying the assumptions listed above. Let $f_{\Delta,h}(\omega)$ and $f_{T,h}(\omega)$ be the marginal discrete spectra of Δ_t and T_t, respectively, at frequency ω for t in block $\mathbf{h}$, and let $\rho_h(\omega)$ and $\theta_h(\omega)$ be their coherence

and phase spectra, so that the spectral matrix at frequency ω, for $\omega = \pi/L, 2\pi/L, ..., \pi$, is

$$f_h(\omega) = \begin{pmatrix} f_{\Delta,h}(\omega) & \rho_h(\omega)\sqrt{f_{T,h}(\omega)f_{\Delta,h}(\omega)}\exp\{i\theta_h(\omega)\} \\ \rho_h(\omega)\sqrt{f_{T,h}(\omega)f_{\Delta,h}(\omega)}\exp\{-i\theta_h(\omega)\} & f_{T,h}(\omega) \end{pmatrix}. \tag{3.1}$$

Then suppose that

$$\begin{cases} f_{\Delta,h}(\omega) = (4\sin^2(\omega/2))^{\gamma_\Delta(X_h)} \times \dfrac{\lambda_\Delta(X_h)}{(\phi_\Delta(X_h) + 4\sin^2(\omega/2))^{\nu_\Delta(X_h)}} \\ f_{T,h}(\omega) = (4\sin^2(\omega/2))^{\gamma_T(X_h)} \times \dfrac{\lambda_T(X_h)}{(\phi_T(X_h) + 4\sin^2(\omega/2))^{\nu_T(X_h)}} \\ \rho_h(\omega) = \dfrac{\lambda_\rho(X_h)}{(\phi_\rho(X_h) + 4\sin^2(\omega/2))^{\nu_\rho(X_h)}} \\ \theta_h(\omega) = \sum_{k=1}^{K} c_k(X_h)\sin(k\omega). \end{cases} \tag{3.2}$$

Here, λ_Δ, λ_T, ϕ_Δ, ϕ_T, and ν_ρ are unknown nonnegative functions; ν_T, ν_Δ, γ_Δ, γ_T, and each c_k are unknown real functions; and K is a known constant. Additionally, $0 \leq \lambda_\rho(x) \leq 1$ for all $x \in [0,1]$ and $\phi_\rho(x) \geq 1$ for all $x \in [0,1]$. These constraints ensure that the above model is valid, although the restrictions on ρ_h are more than necessary but encode the desirable property that coherences be stronger at lower frequencies, which can be observed in the data (shown in Section 3.4).

Fixing temporal block $\mathbf{h}$ and setting $\gamma_T(X_h) = \gamma_\Delta(X_h) = 1$, the model says that $T_t^{(0)}$ and $\Delta_t^{(0)}$ for t in block $\mathbf{h}$ have, marginally, Matérn-like spectra, with a modification for discreteness suggested by Castruccio and Stein (2013). The functions $\lambda_T(X_h)$ and $\lambda_\Delta(X_h)$ describe how the overall level of variability of the processes change with X_h; $\phi_T(X_h)$ and $\phi_\Delta(X_h)$ are inverse-range parameters; and $\nu_T(X_h)$ and $\nu_\Delta(X_h)$ describe the rate of decay of the marginal spectral densities in frequency, essentially smoothness parameters for the discrete-time process. The functions $\gamma_T(X_h)$ and $\gamma_\Delta(X_h)$ allow for additional flexibility in describing the marginal spectral behavior near $\omega = 0$ (unless $\phi = 0$). The extra flexibility

allows for a mid-frequency peak in the spectrum (or not), or a spectrum that approaches infinity as frequency approaches zero.

A Matérn-like form is also used to describe the coherence, similar to the coherence model given in Castruccio and Stein (2013). This model assumes that the coherence is monotonically decreasing in frequency. It is common for coherences to be lower at higher frequencies for environmental processes, and more importantly, the assumption appears to be reasonable upon examining the data.

The phase spectra are parameterized as sums of sine functions, following Stein (2005), to enforce the condition that the discrete phase spectrum be an odd function, and enforcing that it be zero at $\omega = 0$ and π, where the discrete Fourier transform is real. A positive, linear phase relationship in continuous time would indicate that Δ_t leads T_t, and the opposite for a negative, linear relationship. If K is taken to be large, this model is in principle quite general, although the specific model we propose, below, is more restrictive.

In our view, the most substantial assumption underlying model (3.2) is that the spectral matrix within a given hourly block, $\mathbf{h}$, is explained by the covariates X_h. As discussed in Section 3.2, we have found that the level of relative humidity, solar radiation, and fluctuations in wind direction are important covariates that describe changes in the spectral properties of these processes. Let R_t be the level of relative humidity (between 0 and 1) at time t, S_t be the solar radiation (in watts per square meter), and W_t be the wind direction (in radians). Write

$$\bar{R}_h = \frac{1}{L} \sum_{t \in \mathbf{h}} R_t$$

for the average level of relative humidity within block $\mathbf{h}$,

$$\bar{S}_h = \mathbf{1} \left\{ \frac{1}{L} \sum_{t \in \mathbf{h}} S_t > 0 \right\}$$

for the indicator that the sun is up in time block $\mathbf{h}$, and

$$\bar{W}_h = \frac{1}{L}\sum_{t\in\mathbf{h}} \left| e^{iW_{t+1}} - e^{iW_t} \right|$$

for the average modulus of the minute-to-minute changes in wind direction within block $\mathbf{h}$. Let $L = 60$ minutes so that $\mathbf{h}$ is a one-hour block. After considerable experimentation, we have found the following parameterization of model (3.2) in terms of $\bar{R}_h$, $\bar{S}_h$, and $\bar{W}_h$ to be useful for describing the process of interest:

$$\begin{cases} f_{\Delta,h}(\omega) = 4\sin^2(\omega/2) \times \dfrac{\exp\{\lambda_0^\Delta + \lambda_{11}^\Delta \bar{S}_h\}(\bar{W}_h + \lambda_{22}^\Delta)^{\lambda_{21}^\Delta}(1-\bar{R}_h)^{\lambda_{31}^\Delta}}{(\phi_0^\Delta + 4\sin^2(\omega/2))^{\nu_0^\Delta + \nu_1^\Delta \bar{S}_h}} \\ f_{T,h}(\omega) = (4\sin^2(\omega/2))^{\gamma_0^T} \times \dfrac{\exp\{\lambda_0^T + \lambda_{11}^T \bar{S}_h\}(\bar{W}_h + \lambda_{22}^T)^{\lambda_{21}^T}(1-\bar{R}_h + \lambda_{32}^T)^{\lambda_{31}^T}}{(\phi_0^\Delta + 4\sin^2(\omega/2))^{\nu_0^\Delta + \nu_1^\Delta \bar{S}_h}} \\ \rho_h(\omega) = \dfrac{(1-\bar{R}_h)^{\lambda_0^\rho}}{(\phi_0^\rho + 4\sin^2(\omega/2))^{\nu_0^\rho}} \\ \theta_h(\omega) = \displaystyle\sum_{k=1}^{3} \{c_{0k} + c_{1k}\bar{S}_h + c_{2k}\mathbf{1}_{\{\bar{W}_h \geq 0.05\}} + c_{3k}\mathbf{1}_{\{\bar{R}_h < 0.85\}}\} \sin(k\omega), \end{cases} \tag{3.3}$$

where $\lambda_{22}^\Delta, \lambda_{22}^T, \lambda_{32}^T, \phi_0^\Delta, \phi_0^T, \lambda_0^\rho$, and ν_0^ρ are all positive, ϕ_0^ρ is greater than one, and all other parameters are real.

One of the important features of this process is that dew point is bounded above by temperature. We ensure that $\lambda_\Delta(X_h) \downarrow 0$ as $\bar{R}_h \uparrow 1$ in such a way that $f_{\Delta,h}(\omega) \downarrow 0$ as $\bar{R}_h \uparrow 1$ everywhere except possibly when $\omega = 0$, encoding the property that Δ_t does not vary when $\bar{R}_h$ is 100% and therefore that temperature can fall only mildly below dew point. We have found in conditional simulations that the constraint is nearly always satisfied and only mildly disobeyed when so; we provide more details in the Appendix 3.6.2. Additionally, we set $\gamma_\Delta(X_h) = 1$ since, as we will discuss in Section 3.4, ϕ_0^Δ is estimated as essentially zero, and under this condition $\gamma_\Delta(X_h)$ is not estimable.

We have found that this model describes the data well and was chosen with that goal in mind rather than out of more general concerns. The model was chosen to a certain extent based on comparisons with empirical estimates of the components of the spectral matrix, which are displayed alongside the fitted spectra in Figs 3.3 - 3.6, discussed in more detail in Section 3.4. Some rationale for our choices is as follows. The functions λ_Δ and λ_T describe the overall variability of the two processes, and it is natural to parameterize variances as exponential functions. Our parameterization would be the exponential of a linear model in $\bar{S}_h$, $\log(1-\bar{R}_h)$, and $\log(\bar{W}_h)$ if we set λ_{22}^{Δ}, λ_{22}^{T}, and λ_{32}^{T} to zero, but these parameters allow that the variability be nonzero when the covariates approach their boundaries (while we insist that the variability of Δ_t goes to zero as $\bar{R}_h$ approaches 100%). A similar form is used for λ_ρ, but only $\bar{R}_h$ was found to explain the changes in the coherences across hourly blocks. Allowing ν_Δ and ν_T to vary with $\bar{R}_h$ and $\bar{W}_h$ is not a substantial improvement, and neither is allowing for non-constant ν_ρ. The ϕ's are assumed to be constant for the same reason. Finally, the threshold levels used in the model for θ_h were chosen upon examining the data. More details are provided in Section 3.4.

More complicated models were considered and found not to be substantial improvements. In addition to those from models discussed above, the model fits are not much improved by letting the spectral matrix vary with the average level of solar radiation rather than simply an indicator of whether the sun is up. While the actual level of solar radiation probably has a role in describing how the spectral matrix changes, perhaps more data is needed to quantify its effect, which may be hard to distinguish from that of relative humidity. Other model forms considered include different parameterizations of λ_T and λ_Δ, different threshold levels for the functions in θ_h, and sub-models of this stated model.

3.3.2 Likelihood approximation

From a computational point of view, models like this are advantageous because they can be fitted relatively quickly via approximate likelihood methods. As in Section 2.4.1, we use an

independent blocks Whittle likelihood, which treats residual processes as independent across hours makes the approximation that the multivariate periodogram is independent at distinct Fourier frequencies.

Let $F_h(\omega)$ be the discrete Fourier transform of (Δ_t, T_t) in hourly block $\mathbf{h}$ at frequency ω, and let $f_h(\omega)$ be the spectral matrix in temporal block $\mathbf{h}$ at frequency ω, given in (3.1). Then an approximation to the log-likelihood for model (3.3) is

$$l(F_1, ..., F_H; f_1, ..., f_H) \approx -\sum_{\omega,h} \{\log |f_h(\omega)| + \mathrm{tr}(f_h^{-1}(\omega) F_h(\omega) F_h(\omega)^H)\}.$$

Maximizing the approximate likelihood yields (approximately) unbiased estimating equations (to the extent that the periodogram is approximately unbiased), so while the block independence approximation involves some loss of efficiency, as discussed in the next section, it does not induce bias.

As discussed in Section 2.4.1, there are other approximate likelihood methods for locally stationary processes that may yield better approximations at some additional computational expense. One of the benefits of the approximation proposed by Guinness and Stein (2013b) is that their approximation accounts for between-block correlation of the process, whereas the independent blocks likelihood does not. Since our interest in this chapter is specifically in the high-frequency behavior of the process, we view our approach as a substantial and beneficial simplification. A continuous time model would present additional modeling challenges not central to the goals of this chapter – that is, characterizing whatever low-frequency power remains in the residual process after removing the smoothed component – and would rely on more computationally burdensome methods for model fitting. Our finding that the fitted spectra are not too sensitive to the choice of block length, discussed above and in more detail in the Appendix 3.6.1, is an indication that our approach is reasonable for estimating the quantities of interest.

3.3.3 Standard errors of estimates

We employ a block bootstrapping scheme for estimating the standard errors of the estimates of model (3.3). We use Carlstein's rule (Carlstein, 1986) for the blocks, following the terminology of Hall et al. (1995) – that is, non-overlapping blocks – where our procedure is, for bootstrap samples $k = 1, ..., K$,

1. Sample a vector $s^{(k)} \sim \text{Multinom}(48; 1/48, ..., 1/48)$ of disjoint days with replacement, independently of $s^{(k-1)}, ..., s^{(1)}$.

2. Let $\Delta^{(k)} = (\Delta_{\mathbf{s_1^{(k)}}}, ..., \Delta_{\mathbf{s_{48}^{(k)}}})$ (i.e. the difference residuals in each day of $\mathbf{s^{(k)}}$), and similarly define $T^{(k)}$, $\bar{R}^{(k)}$, $\bar{W}^{(k)}$, and $\bar{S}^{(k)}$.

3. Fit model (3.3) on $(\Delta^{(k)}, T^{(k)})$ given $\bar{R}^{(k)}$, $\bar{W}^{(k)}$, and $\bar{S}^{(k)}$.

Then taking standard deviations gives estimates of marginal standard errors of the parameter estimates. The blocks are chosen as day-long in size under the assumption that this gives a bootstrap sample that more accurately follows the distribution of (Δ_t, T_t), given the nonstationarity of the data and the diurnal cycles of the processes. For our analysis, we take the number of bootstrap samples to be $K = 100$.

3.4 Data Analysis and Results

We explore the results of fitting model (3.3) to the data described in Section 3.2. The analysis has been carried out on the full model in addition to some sub-models, and we have found that the model corresponding to letting T_t be, marginally, uniformly modulated, is as adequate as the full model based on the criterion discussed below.

3.4.1 A criterion for likelihood comparisons

For problems with many observations, likelihood ratio tests will very often be significant. In such settings, and when all of the models in question are most likely wrong, it is often more

relevant to ask how much worse one model is than another rather than simply whether one model is worse than another at all.

If the log-likelihood ratio between the true model and a proposed model scales like the number of observations times the Kullback-Leibler (KL) divergence between the two models, as in the i.i.d. setting, then dividing the log-likelihood ratio by the number of observations would seem to be a more informative measure of the distance between the two models, if the goal is informal model comparison rather than testing. For discrete-time Gaussian time series with square-integrable log-spectra, Amari and Nagaoka (2007) shows that the expected log-likelihood ratio (that is, the KL divergence) for a finitely observed time series scales like the number of observations times the KL divergence between the two spectral densities, under increasing-domain asymptotics. While other model selection criteria, like the Bayesian Information Criterion (BIC), include penalties for the number of observations, these still treat the parameter count as the object of interest; however, there are often other important informal considerations beyond the parameter count, such as the simplicity of the model form or the number of covariates in the model. We suggest that if the distance between two models is not substantially large, we should choose the most palatable model, however we define "palatable" for the problem at hand. For the process we are modeling, we believe the most palatable models are those that describe the process in a parsimonious and interpretable way without imposing unnatural restrictions and while still accounting for the important features observed in the data.

Table 3.1 displays the difference between the approximate log-likelihoods of the full model (3.3) and sub-models of interest, divided by the number of observations. As a suggested rule of thumb, differences bigger than 0.1 may be considered large and differences bigger than 0.01 may be considered potentially important, whereas those less than 0.01 may be regarded as not substantially large. We have found this rule of thumb to be adequate based on experience, but are not claiming that this rule is optimal under any theoretical framework.

The full model can be simplified without much consequence to the approximate likelihood, under the above rule, by letting T_t be, marginally, uniformly modulated, which corresponds to setting $\nu_1^T = 0$. By uniform modulation we mean that one or both of the processes can be described (marginally or jointly) as the product of a stationary process and a function varying in time (Priestley, 1981). The usual likelihood ratio statistic comparing this reduced model to the full model is about 104 on one degree of freedom. While this is only an approximate likelihood ratio, this suggests that the usual p-value corresponding to the null hypothesis that the reduced model is true would be extremely small; that is, the distance between the two models is nonzero. However, by comparison, the difference in approximate log-likelihood per observation between this reduced model and the full model is about 0.002, which we consider small enough to be insubstantial. While the BIC for the reduced model is also larger than that for the full model, we view uniform modulation as substantially simpler than, for example, removing one parameter from λ_T, even though the penalty in the BIC would be the same in each case. For the remainder of this discussion, we discuss the reduced model.

We find that the reduced model is better than other compared sub-models from the perspective of our criterion. The largest gains in likelihood come from letting $\lambda_T(X_h)$ and $\lambda_\Delta(X_h)$ vary with both $\bar{R}_h$ and $\bar{W}_h$ rather than with only one of the two. The contribution of $\bar{S}_h$ is substantial but much less so than that of either $\bar{R}_h$ or $\bar{W}_h$.

Allowing for a non-zero phase also contributes substantially to the likelihood, as does letting the phase vary with the covariates, with an indication that the temperature residual often leads the difference (and therefore dew point) residual and that this relationship might reverse when relative humidity is low but the sun is also down, discussed below.

3.4.2 Features of fitted model

Table 3.2 displays the fitted parameter values for the reduced model, setting $\nu_1^T = 0$, corresponding to letting T_t be marginally uniformly modulated. Overall variability of both

Model Name	Params Excluded	# Params	All Days (69,120 obs) log-Lik Diff / # obs
Full Model	–	33	0
Uniformly modulated T_t †	ν_1^T	32	0.001
ν_T and ν_Δ constant	ν_1^T, ν_1^Δ	31	0.031
Joint uniform modulation of (Δ_t, T_t)	$\nu_1^T, \nu_1^\Delta, \lambda_0^\rho, c_{1k}, c_{2k}, c_{3k}$	21	0.101
f_T and f_Δ don't depend on $\bar{S}_h$	$\lambda_{11}^T, \lambda_{11}^\Delta, \nu_1^T, \nu_1^\Delta$	29	0.067
f_T and f_Δ don't depend on $\bar{W}_h$	$\lambda_{21}^T, \lambda_{22}^T, \lambda_{21}^\Delta, \lambda_{22}^\Delta$	29	0.175
f_T and f_Δ don't depend on $\bar{R}_h$	$\lambda_{31}^T, \lambda_{32}^T, \lambda_{31}^\Delta$	30	0.434
Constant Phase	c_{1k}, c_{2k}, c_{3k}	24	0.066
Phase = 0	$c_{0k}, c_{1k}, c_{2k}, c_{3k}$	21	0.085

Table 3.1: (Full model approximate likelihood - reduced model approximate likelihood) / number of observations. Dropping any of the three covariates results in substantial losses in approximate likelihood, and there is evidence of a nonzero phase relationship. †We take the model corresponding to T_t being marginally uniformly modulated as the best model. In the ensuing discussion, this model is referred to simply as "the reduced model".

processes is predicted to increase as $\bar{R}_h$ decreases and $\bar{W}_h$ increases. Interestingly, variability of Δ_t is predicted to increase under daylight conditions, but variability of T_t decreases under daylight conditions, holding $\bar{R}_h$ and $\bar{W}_h$ constant. Smoothness of the difference residual decreases under daylight conditions.

Lower $\bar{R}_h$ favors a more negative phase relationship, whereas zero $\bar{S}_h$ and lower $\bar{W}_h$ favor a more positive phase relationship. A negative phase relationship indicates that temperature leads the difference, and the opposite for a positive phase relationship. To the extent that a phase relationship is present, the time lag is on the order of a minute or less. Finally, the decrease in coherence as $\bar{R}_h$ increases is slow over the observed values of $\bar{R}_h$.

While both ϕ_0^Δ and λ_{22}^Δ are estimated as essentially zero, we view their exclusion from the model as unnatural. The former, an inverse range parameter, is estimated as zero because the blocks are only one hour long and our approximation to the likelihood ignores dependence across the blocks. If the latter were exactly zero, this would correspond to the constraint that $\Delta_t = 0$ when $\bar{W}_h = 0$; while it is possible that this is nearly true, we know of no physical argument for why this must be true and so therefore prefer not to force λ_{22}^Δ to be exactly zero.

Some of the bootstrap standard errors of the parameter estimates are moderately large;

Parameter Functional Form		Parameter	Estimate	(s.e.)
$\lambda_T(\bar{X}_h)$	$\exp\{\lambda_0^T + \lambda_{11}^T \bar{S}_h\}(\bar{W}_h + \lambda_{22}^T)^{\lambda_{21}^T}(1 - \bar{R}_h + \lambda_{32}^T)^{\lambda_{31}^T}$	λ_0^T	0.23	(0.37)
		λ_{11}^T	-0.126	(0.003)
		λ_{21}^T	1.68	(0.26)
		λ_{22}^T	0.04	(0.01)
		λ_{31}^T	1.00	(0.35)
		λ_{32}^T	0.09	(0.13)
$\phi_T(\bar{X}_h)$	ϕ_0^T	ϕ_0^T	0.76	(0.16)
$\nu_T(\bar{X}_h)$	$\nu_0^T + \nu_1^T \bar{S}_h$	ν_0^T	1.42	(0.13)
		ν_1^T	–	
$\gamma_T(\bar{X}_h)$	γ_0^T	γ_0^T	-0.01	(0.03)
$\lambda_\Delta(\bar{X}_h)$	$\exp\{\lambda_0^\Delta + \lambda_{11}^\Delta \bar{S}_h\}(\bar{W}_h + \lambda_{22}^\Delta)^{\lambda_{21}^\Delta}(1 - \bar{R}_h)^{\lambda_{31}^\Delta}$	λ_0^Δ	0.58	(0.20)
		λ_{11}^Δ	0.915	(0.003)
		λ_{21}^Δ	0.93	(0.08)
		λ_{22}^Δ	0.00	(0.00)
		λ_{31}^Δ	1.76	(0.09)
$\phi_\Delta(\bar{X}_h)$	ϕ_0^Δ	ϕ_0^Δ	0.003	(0.001)
$\nu_\Delta(\bar{X}_h)$	$\nu_0^\Delta + \nu_1^\Delta \bar{S}_h$	ν_0^Δ	1.29	(0.02)
		ν_1^Δ	-0.441	(0.003)
$\lambda_\rho(\bar{X}_h)$	$(1 - \bar{R}_h)^{\lambda_0^\rho}$	λ_0^ρ	0.11	(0.04)
$\phi_\rho(\bar{X}_h)$	ϕ_0^ρ	ϕ_0^ρ	1.39	(0.07)
$\nu_\rho(\bar{X}_h)$	ν_0^ρ	ν_0^ρ	0.54	(0.05)
$c_k(\bar{X}_h)$	c_{0k}	c_{01}	-0.81	(0.26)
		c_{02}	0.07	(0.03)
		c_{03}	-0.20	(0.11)
	$c_{1k}\bar{S}_h$	c_{11}	-1.85	(0.26)
		c_{12}	0.37	(0.03)
		c_{13}	-0.32	(0.11)
	$c_{2k}\mathbf{1}_{\{\bar{W}_h \geq 0.05\}}$	c_{21}	-0.35	(0.48)
		c_{22}	-0.10	(0.06)
		c_{23}	-0.01	(0.23)
	$c_{3k}\mathbf{1}_{\{\bar{R}_h < 0.85\}}$	c_{31}	1.52	(0.12)
		c_{32}	-0.21	(0.07)
		c_{33}	0.31	(0.06)

Table 3.2: Parameter estimates for the reduced version of model (3.3), setting $\nu_{T1} = 0$. Overall variability is predicted to increase as $\bar{R}_h$ decreases and $\bar{W}_h$ increases. Variability of Δ_t is predicted to increase under daylight conditions, but variability of T_t decreases under daylight conditions. Smoothness of the difference residual decreases in daylight. Coherence decreases slowly as $\bar{R}_h$ increases. The phase relationship is usually negative – T_t leads Δ_t – more so when the sun is up, $\bar{R}_h$ is low, and $\bar{W}_h$ is large. A positive phase relationship – Δ_t leads T_t – is predicted when $\bar{R}_h$ is low but $\bar{W}_h$ is also low and the sun is down. The time lag of the phase relationship is on the order of a minute or less.

however, some parameter estimates are strongly correlated and there are linear combinations of the parameters that are estimated more precisely. To illustrate, Figure 3.2 displays the marginal and pairwise distributions of the block bootstrap parameter estimates for the parameters in $f_{T,h}$. The parameter estimates that are most strongly correlated are those that share a similar effect on the shape of the marginal spectra; as one example, estimates of λ_{21}^T and λ_{22}^T are positively correlated since both parameters control the degree to which the level of $\bar{W}_h$ influences the overall variability of T_t. For similar reasons, λ_{31}^T and λ_{32}^T are strongly correlated, as are ϕ_0^T and ν_0^T. Similar relationships are present in the parameter estimates in the other components of the spectral matrix. While some parameters may be set equal to zero with little consequence to the likelihood, we choose our parameterization in part for its interpretability. The implication of this, however, is that there are linear combinations of the parameters in the model that are better estimated than the parameters themselves, and therefore that the groups of parameters with larger marginal standard errors should be interpreted jointly.

Figs 3.3, 3.4, 3.5, and 3.6 display the average fitted marginal spectra, coherence, and phase spectra from the reduced model fitted to all of the data, along with empirical estimates from the multivariate periodogram, where in each of the figures we average the marginal, co- and quadrature- spectra over bins corresponding to the level of either $\bar{R}_h$, $\bar{W}_h$, the hourly average level of solar radiation, or all three, respectively, to obtain smoothed estimates. We have not smoothed across frequencies. Overall, the model appears to track the changes in the spectral behavior of the processes well. To the extent that there are interactions between the covariates, our model is largely able to cope with them. In general the fit is better for the marginal spectra than for the phase or coherence. For the phase in particular, there is empirical evidence that the phase relationship changes from temperature leading the difference to the difference leading temperature as solar radiation and wind direction vary, but this is somewhat difficult to fully capture in the model, even with the presence of these covariates in the phase model.

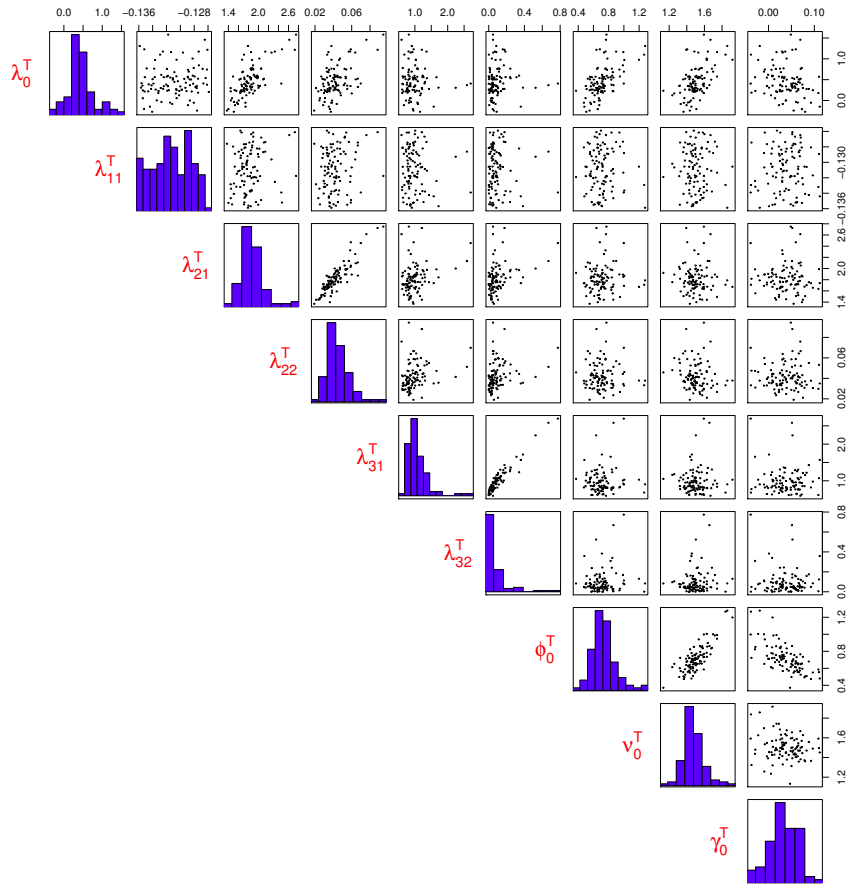

Figure 3.2: Estimates of marginal and pairwise distributions of block bootstrap parameter estimates for $f_{T,h}$. The model form is given in (3.3). The parameters that are most strongly correlated are those that share a similar influence on the shape of the marginal spectrum.

Although the correlation between the three covariates prevents us from drawing strong conclusions about the covariates' effects on the process, the prediction that volatility of the temperature residual decreases in daylight conditions after holding the other covariates constant suggests that our model is able to differentiate the effects of the three covariates. Our model captures the broad changes in the spectral behavior of the process by letting the overall variability change with the three covariates and letting the smoothness of the difference residual vary only with the indicator that the sun is up. While each covariate may have an important effect on the overall behavior of the process, their effects may be more concentrated on different aspects of the overall behavior.

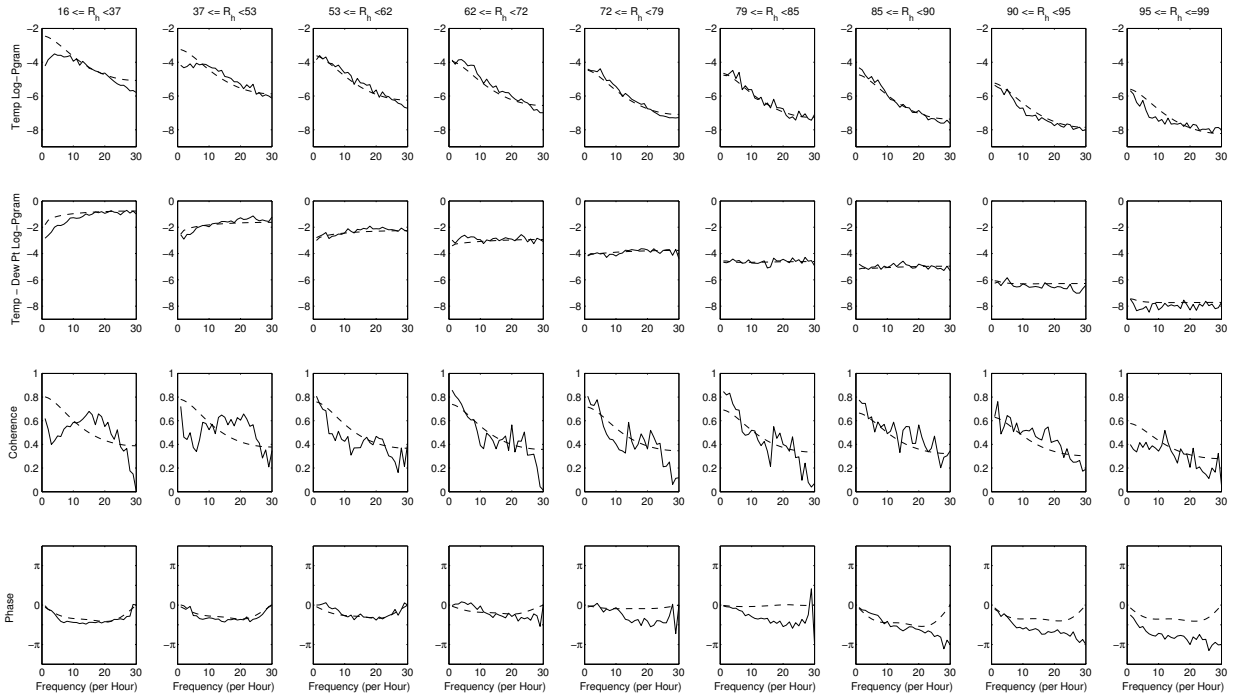

Figure 3.3: Fitted (dashed) and empirical (solid) marginal spectra, coherence, and phase, averaged across hours binned by the level of $\bar{R}_h$. The bins were chosen to contain approximately equal numbers of hours. The empirical estimates change as $\bar{R}_h$ increases, with the changes mostly tracked by the fitted spectra. The difference residual appears to become less variable and apparently $\nu_\Delta(X_h)$ increases as $\bar{R}_h$ increases, and the temperature residual becomes less variable as $\bar{R}_h$ increases. Since $\nu_\Delta(X_h)$ does not depend on relative humidity, any apparent changes in the smoothness of the difference residual is evidently tracked by the proportion of the time the relative humidity is at a given level during the day versus night. Coherence appears to decrease as $\bar{R}_h$ increases, tracked by the model, and there is an apparent phase relationship with T_t leading Δ_t, especially for large $\bar{R}_h$, somewhat tracked by the model.

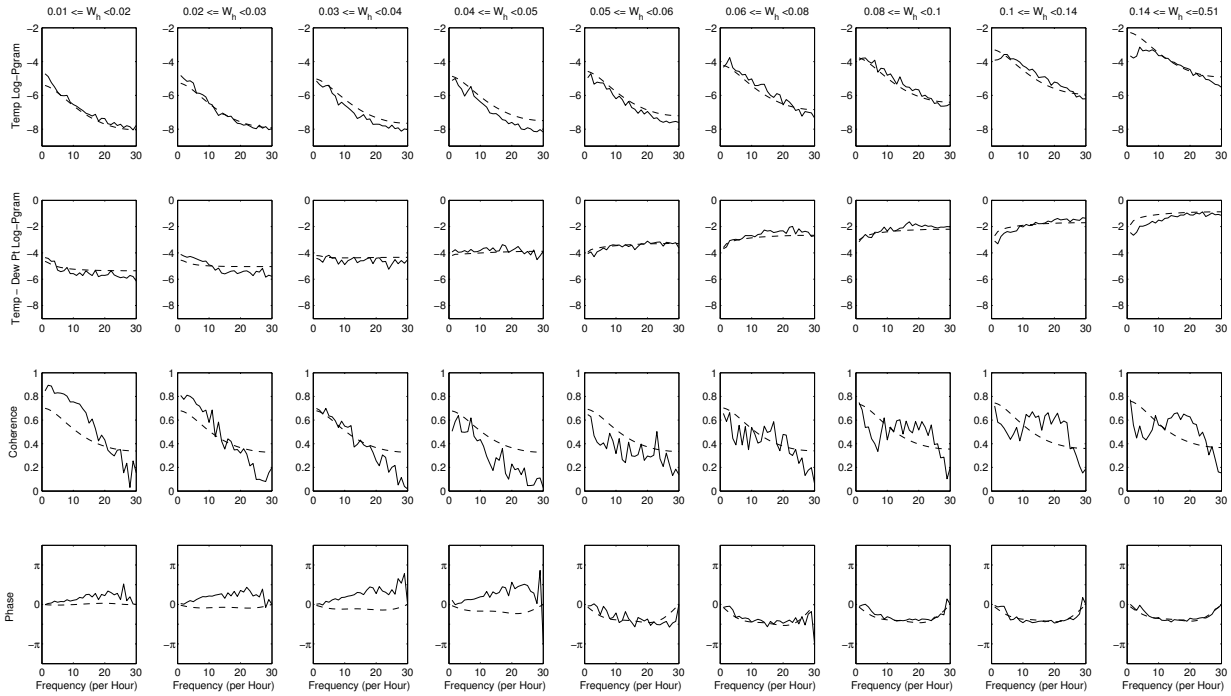

Figure 3.4: Fitted (dashed) and empirical (solid) marginal spectra, coherence, and phase, averaged across hours binned by the level of $\bar{W}_h$. The bins were chosen to contain approximately equal numbers of hours. The empirical estimates change as $\bar{W}_h$ increases, with the changes mostly tracked by the fitted spectra. The difference residual becomes more variable and apparently $\nu_{\Delta}(X_h)$ decreases as $\bar{W}_h$ increases, and the temperature residual becomes more variable as $\bar{W}_h$ increases. Since $\nu_{\Delta}(X_h)$ does not depend on wind direction, any apparent changes in the smoothness of the difference residual are evidently tracked by the proportion of the time the magnitude of changes in wind direction is at a given level during the day versus night. Coherence appears relatively constant in $\bar{W}_h$. There is an apparent reversal in phase direction around where $\bar{W}_h = 0.05$, with Δ_t leading T_t for small $\bar{W}_h$ and the reverse for large $\bar{W}_h$. This reversal is not fully captured by the model.

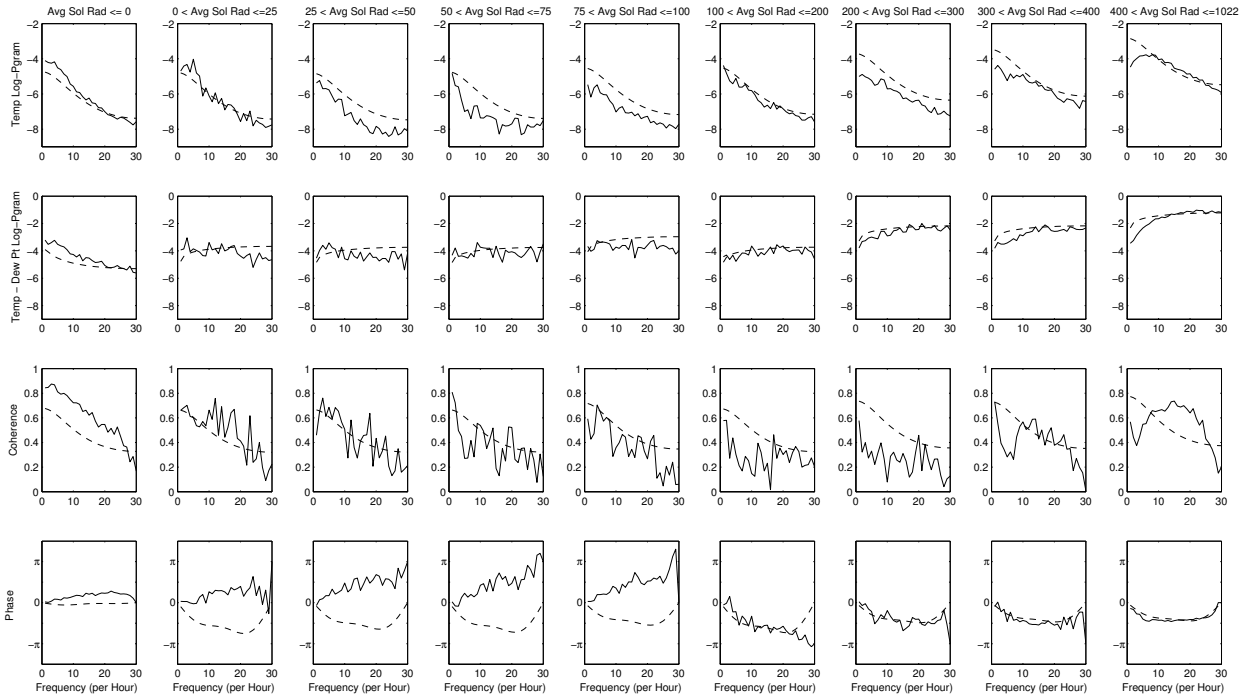

Figure 3.5: Fitted (dashed) and empirical (solid) marginal spectra, coherence, and phase, averaged across hours binned by the average level of solar radiation. The first and last bins contain about 44% and 25% of the hours, respectively, and the middle seven each contain approximately 4% of the hours. The empirical estimates change as solar radiation increases, with the changes mostly tracked by the fitted spectra. The difference residual becomes somewhat more variable and apparently $\nu_\Delta(X_h)$ decreases as radiation increases, with the major distinction being between daytime and nighttime hours. The temperature residual becomes somewhat less variable in daytime versus nighttime hours. Coherence appears relatively constant in $\bar{S}_h$. There is an apparent reversal in phase direction when the average level of solar radiation is small but positive, with Δ_t leading T_t for low solar radiation and the reverse for high solar radiation. This reversal is not captured completely by the model.

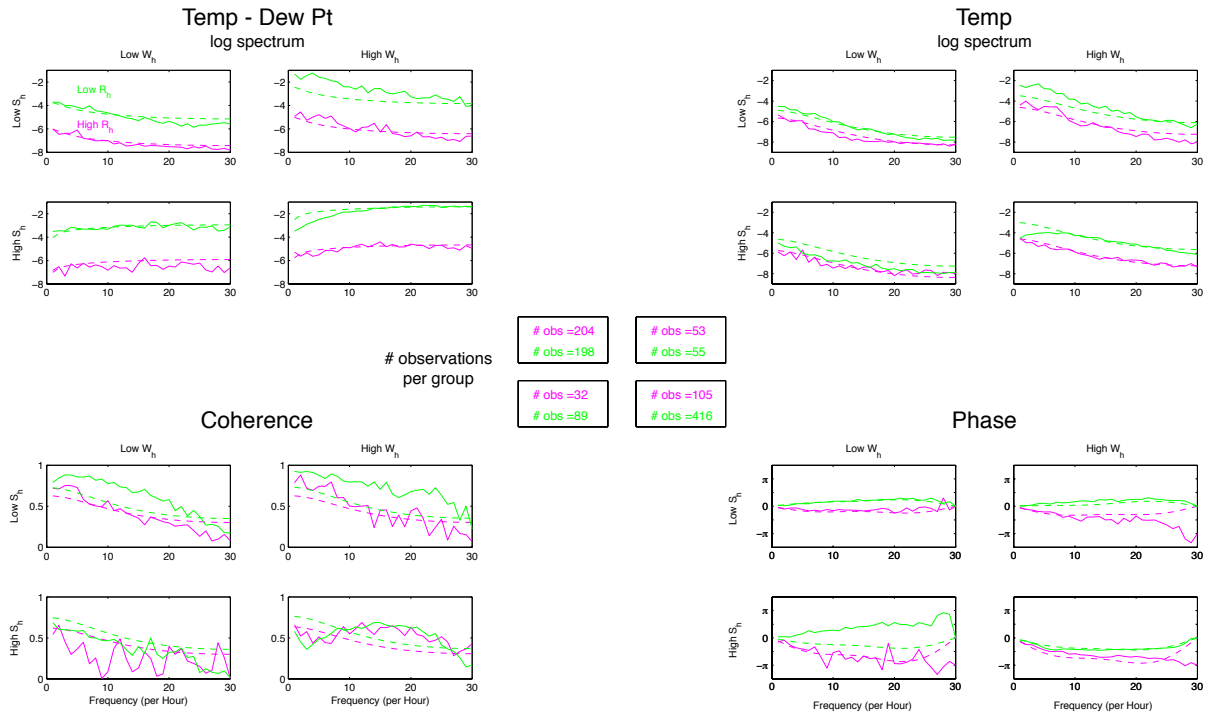

Figure 3.6: Interactions between covariates. We display fitted (dashed) and empirical (solid) marginal spectra (top left and right), coherence (bottom left) and phase (bottom right), binned by levels of $\bar{R}_h$, $\bar{W}_h$, and $\bar{S}_h$. Bins are identified by the axis labels and color, with green corresponding to low $\bar{R}_h$ and magenta to high $\bar{R}_h$. The center image gives the numbers of observations in each of the eight bins. $\bar{R}_h$ levels are $> 85\%$ (high) and $\leq 85\%$ (low); $\bar{W}_h$ levels are ≥ 0.05 (low) and < 0.05 (high); $\bar{S}_h$ levels are $= 1$ (high) and $= 0$ (low). There is some systematic misfit present, perhaps most notably when $\bar{R}_h$ is low and $\bar{S}_h$ is high but $\bar{W}_h$ is low. In that scenario, we underestimate the overall level of variability in T_t and overestimate it in Δ_t, and apparently cannot predict the direction of the phase relationship. But to the extent that there are interactions between the three covariates, our model is apparently mostly able to account for them.

3.5 Discussion

We have provided a model for the bivariate behavior of the temporal high-frequency changes in temperature and dew point, conditionally on the level of relative humidity, the magnitude of the minute-to-minute changes in wind directions, and the presence or lack of sunlight. We are able to capture the changes in the behavior of this bivariate process over time and our model provides an interpretable description of those changes in terms of the covariates. Importantly, our model successfully copes with the physical constraint of the process, that dew point is bounded above by temperature.

Our model does impose some strong assumptions on the process, namely that that the process can be explained entirely by the three included covariates. While not explored in detail here, model (3.3) can in principle be extended to relax this assumption. Estimates from our model will be reasonable if this assumption is, loosely speaking, on average true. Indeed, as shown in Section 3.4, our model seems to perform well on average; however, we have observed that the ratio of fitted to empirical spectra is over-dispersed when compared across hours, an indication that there are changes in variability in the processes that are unexplained by the covariates included in the model. As an extension, one might introduce random effects for each block $\mathbf{h}$ to allow there to be deviations from the proposed model in each hourly block. However, since our goal is explaining the changes in volatility with known covariates, we are more interested in identifying other important covariates than in modeling unexplained variation by a random effect. Such a modification would also come at increased computational costs.

The modeling of fine timescale meteorology may inform physical models of larger scale processes. This chapter represents what we believe is a step in addressing the challenging statistical problem of adequately characterizing the rich data available for these fine timescale processes. Those interested in modeling all of the measured quantities relevant to this system, and in spatial interpolation of the whole multivariate process, will encounter new and interesting statistical and computational challenges relevant to characterizing the complex

behavior we observe in high-frequency meteorological data.

3.6 Appendix

3.6.1 Sensitivity Analysis

We investigate the sensitivity of our analysis to two choices: the length of the window used to estimate the smoothed and residual components of the data, described in Section 3.2; and the length, L, of the blocks in which (T_t, Δ_t) are assumed to be stationary, described in Section 3.3. Performing the same analysis described in the main text, we vary the length of the smoothing window between 1 hour, 2 hours, 3 hours, and 4 hours, and similarly the block length between 30 minutes, 1 hour, 1.5 hours, and 2 hours. We estimate the model parameters in the reduced version of model (3), taking $\nu_1^T = 0$, using each combination of the two choices, for a total of 16 sets of estimates. For both choices, there is of course a tradeoff between lengths that are either longer or shorter. If the block length is chosen to be too short, the periodogram is a more biased estimator of the spectrum, whereas if the block length is too long the approximation of stationarity is less likely to be reasonable. Similarly, smoothing windows that are too long will leave too much of the low frequency behavior of the process in the residual component, whereas windows that are too short will remove high frequency behavior from the residual component, which is the behavior of interest in this chapter.

Table 3.3 displays the average estimates for each choice of smoothing parameter or block length, where we average the four estimates obtained under the same choice of smoothing window length or block length. The parameter estimates do not appear to be very sensitive to either of these two choices, at least within reasonable ranges of choices. Where there are moderate differences, however, the differences do appear to be systematic. For example, the estimate of ϕ_0^T, the inverse range parameter for T_t, increases with the length of the smoothing window and decreases with the length of the time block. The opposite is true of

ν_0^T, the "smoothness" parameter for T_t.

While such systematic differences in estimates are apparent for a few parameters, the differences are small enough that collectively the effect on the fitted spectra is small. To illustrate, Figs 3.7, 3.8, and 3.9 display the average estimated marginal spectra, phase and coherence for each of the sixteen combinations of smoothing window and time block choice, binned by level of relative humidity, wind direction changes, and solar radiation, respectively, as in Figs 3, 4, and 5. The fitted components of the spectral matrix are quite similar on average for each of the sixteen choices, with larger differences occurring on average at the low frequencies, which are not of primary interest to us. Figure 3.6.1 provides a close-up of low frequency part of Figure 3.7.

The biggest differences in the fits appear to be when comparing across blocks by level of solar radiation when solar radiation is nonzero but still moderately low. This is not surprising, as these blocks typically correspond to times in the morning or afternoon when the sun is rising or setting, and therefore if the block length is chosen to be too large, stationarity is unlikely to hold. This suggests that choosing a block length that is reasonably short is advisable, which is why we take the blocks to be one hour in length.

3.6.2 Conditional Simulations

One of the important features of the process we are modeling is that dew point is bounded above by temperature. Our model copes with this constraint by insisting that the variability of Δ_t goes to zero as relative humidity approaches 100%. While this still allows that temperature can fall mildly below dew point, we have found from conditional simulations that this constraint is very rarely violated and is only violated very mildly when so. To illustrate, we provide a conditional simulation of the process given the values of $\bar{R}_h$, $\bar{W}_h$, $\bar{S}_h$, and the smoothed components of temperature and dew point observed in the data. Figure 3.11 displays the first two days of one of these conditional simulations.

Given the values of $\bar{R}_h$, $\bar{W}_h$, and $\bar{S}_h$, let $\hat{f}_h(\omega)$ be the estimated spectral matrix within

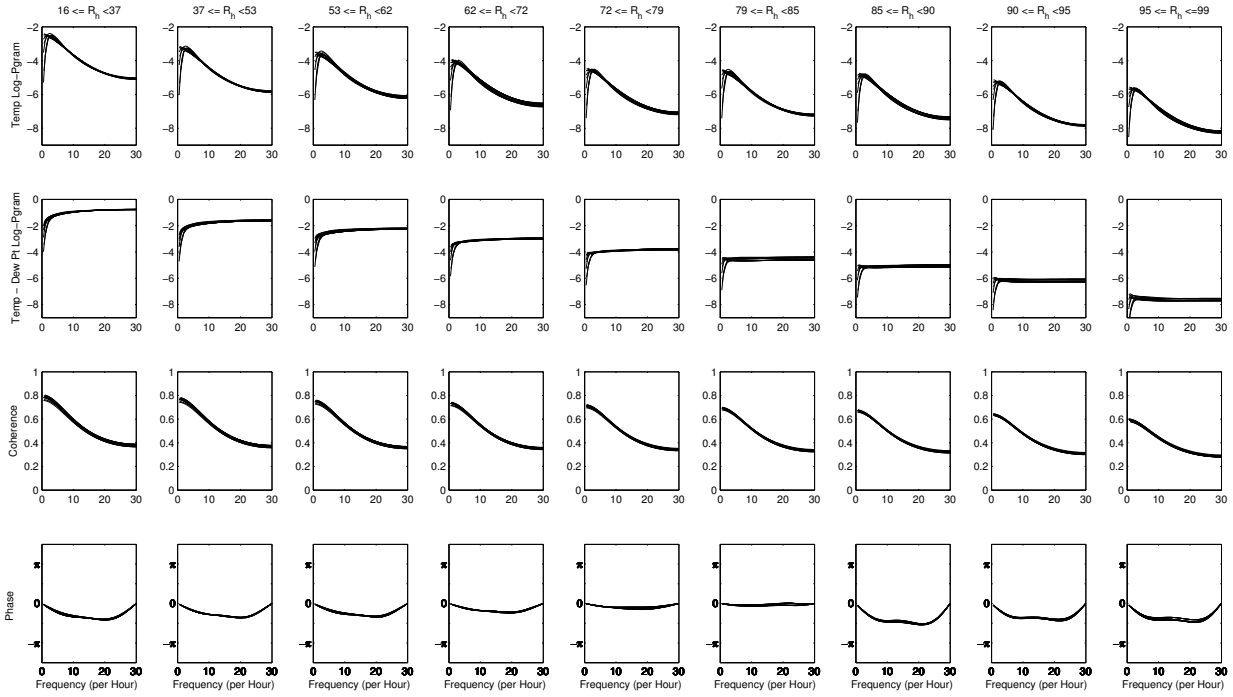

Figure 3.7: Average fitted marginal spectra, coherence, and phase, averaged across blocks binned by the average level of relative humidity. Bins are the same as in Figure 3.3. Each of the sixteen curves in a given plot corresponds to estimates obtained from one choice of smoothing window length and block length. The effect of these choices on the fitted spectra is on average small, with the biggest differences in the estimates at the low frequency, which are not of primary interest to us.

time block $\mathbf{h}$ and at frequency ω, using the reduced version of model (3) (i.e. taking $\nu_1^T = 0$).

Let

$$Z_j \sim \begin{cases} N(0, \hat{f}_h(2\pi j/60)), & j = 0, 30 \\ CN(0, \hat{f}_h(2\pi j/60)), & \text{otherwise} \end{cases}$$

with Z_j independent of $Z_{j'}$ if $|j| \neq |j'|$ and $Z_j = Z_{-j}^*$, where x^* is the complex conjugate of x. Then a conditional simulation of (Δ_t, T_t) is given by $(\tilde{\Delta}_t, \tilde{T}_t)$, where

$$(\tilde{\Delta}_t, \tilde{T}_t) = \sqrt{60} \sum_{j=-29}^{30} Z_j \exp\{2\pi i j t/60\}.$$

We un-difference $(\tilde{\Delta}_t, \tilde{T}_t)$ to produce $(\tilde{\Delta}_t^{(0)}, \tilde{T}_t^{(0)})$, where for each day we take the initial

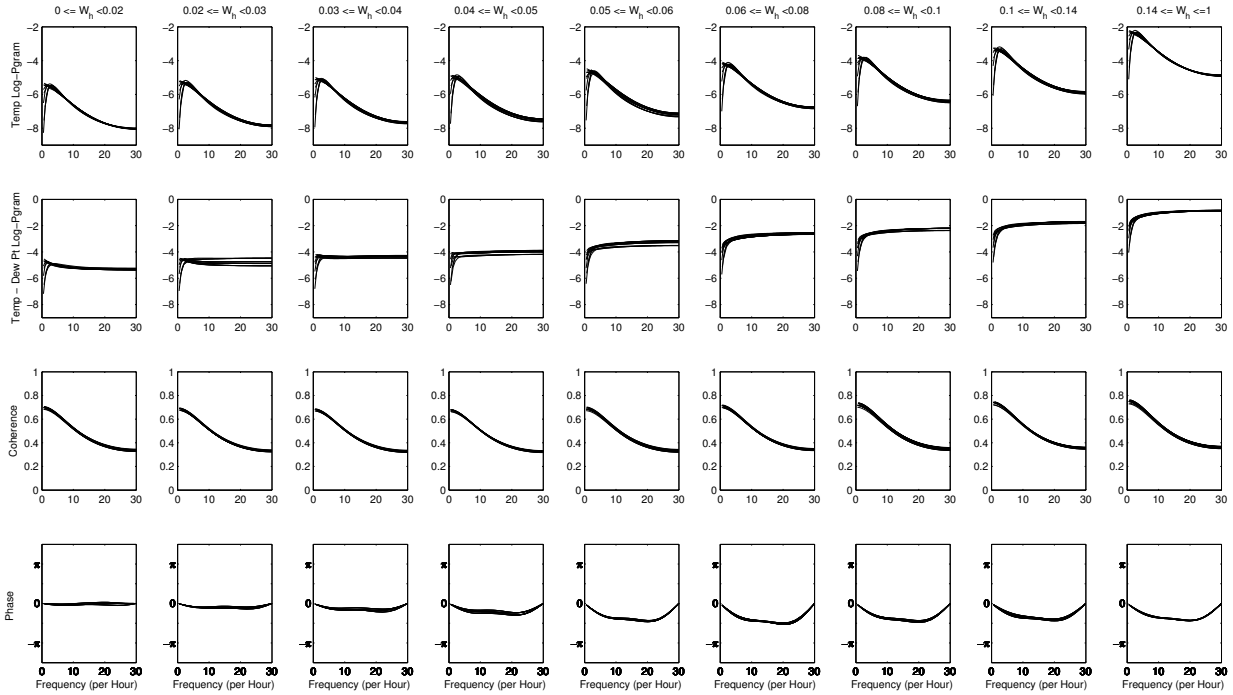

Figure 3.8: Average fitted marginal spectra, coherence, and phase, averaged across blocks binned by the average level of minute to minute changes in wind direction. Bins are the same as in Figure 3.4. Each of the sixteen curves in a given plot corresponds to estimates obtained from one choice of smoothing window length and block length. The effect of these choices on the fitted spectra is on average small, with the biggest differences in the estimates at the low frequency, which are not of primary interest to us.

value of the undifferenced process to be that given by the data:

$$(\tilde{\Delta}_t^{(0)}, \tilde{T}_t^{(0)}) = (\Delta_1^{(0)}, T_1^{(0)}) + \sum_{k=1}^{t} (\tilde{\Delta}_k, \tilde{T}_k),$$

and finally we add $(\tilde{\Delta}_t^{(0)}, \tilde{T}_t^{(0)})$ to the smoothed component computed from the data to produce a conditional simulation.

From Figure 3.11, we see that the conditional simulation reproduces the features we observe in the data. In particular, it is nearly always true that temperature is greater than dew point, although there is a short period in the morning of the second day where the constraint is very mildly satisfied. For the simulation we produced, there are only 60

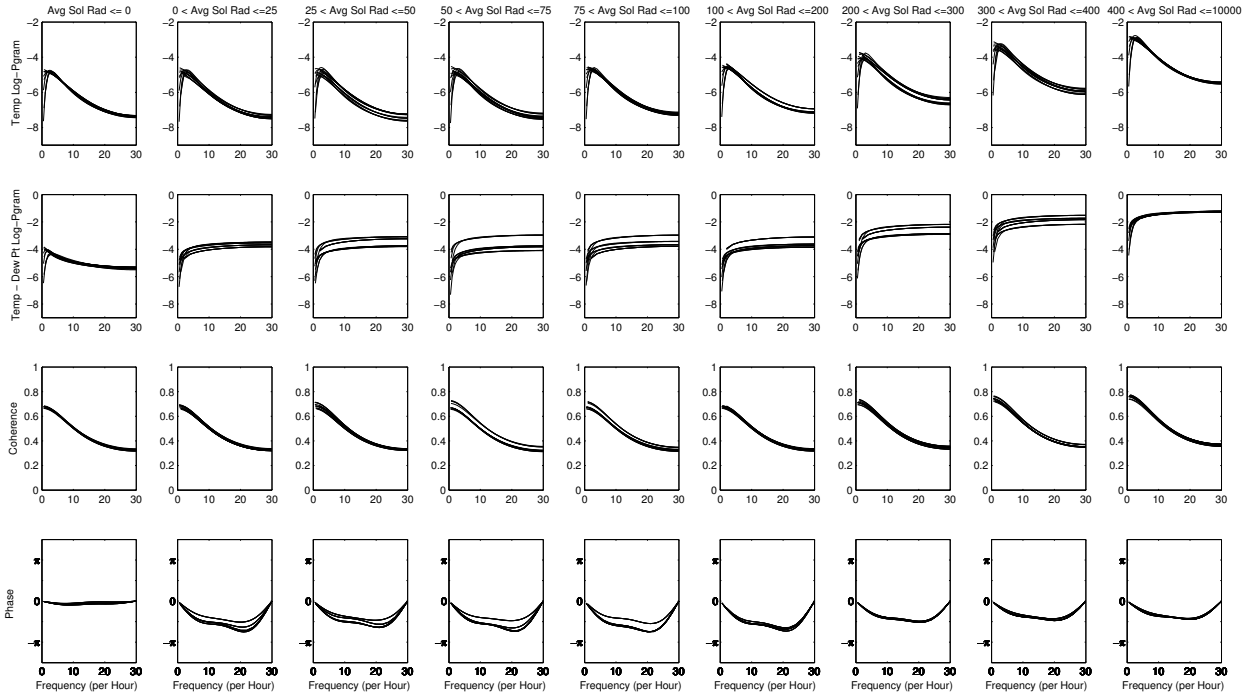

Figure 3.9: Average fitted marginal spectra, coherence, and phase, averaged across blocks binned by the average level of solar radiation. Bins are the same as in Figure 3.5. Each of the sixteen curves in a given plot corresponds to estimates obtained from one choice of smoothing window length and block length. There are some reasonably substantial differences in the fitted marginal spectra during periods when the sun is up but the level of solar radiation is not very high, which typically correspond to periods in the morning or afternoon when the sun is rising or setting. During these periods, the approximation of stationarity is unlikely to hold well if the time window is too long, which suggests picking a reasonably short window.

minutes (out of 69,120) where temperature falls below dew point, and at a maximum falls below dew point by only 0.1 degrees Celsius.

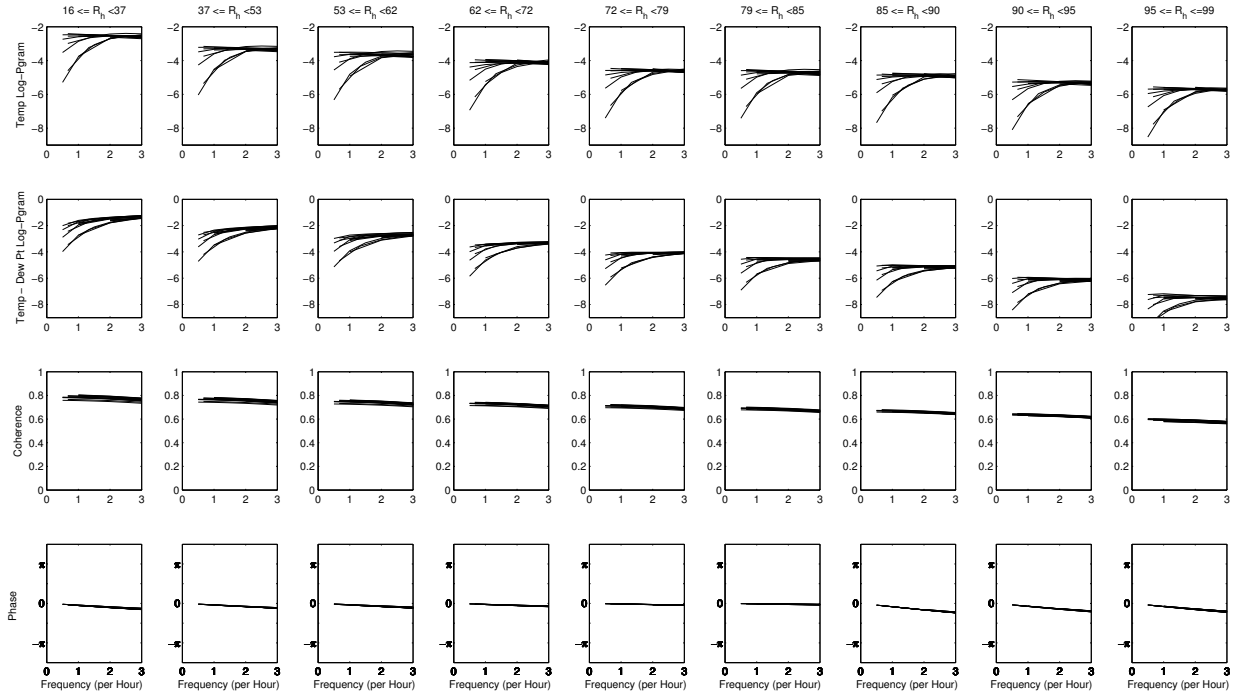

Figure 3.10: Same as Figure 3.7, but restricting attention to the low frequencies. The curves begin at different frequencies because they correspond to different block lengths.

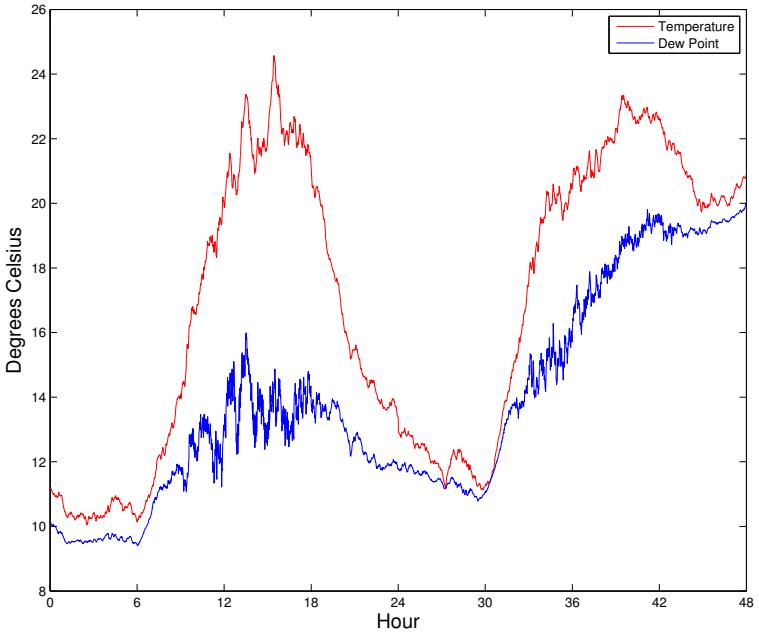

Figure 3.11: Conditional simulation of temperature and dew point, given the average hourly levels of relative humidity, dew point, and minute-to-minute changes in wind direction on May 2 and 3, 2003 of the observed data as well as the computed smoothed components of temperature and dew point on those days; see Section 3.6.2 for more details. The conditional simulation reproduces the important features of the data. In particular, the constraint that dew point is bounded above by temperature is nearly always satisfied, with temperature falling very mildly below dew point in the morning of the second day.

	Smoothing Window				Block Length			
Parameter	1 hour	2 hours	3 hours	4 hours	30 minutes	1 hour	1.5 hours	2 hours
λ_0^T	-0.28	-0.03	0.16	0.26	0.25	-0.00	-0.03	-0.10
λ_{11}^T	-0.15	-0.14	-0.14	-0.14	-0.12	-0.13	-0.17	-0.16
λ_{21}^T	1.64	1.64	1.64	1.64	1.52	1.68	1.67	1.70
λ_{22}^T	0.03	0.03	0.03	0.03	0.03	0.04	0.03	0.04
λ_{31}^T	0.99	0.99	1.00	1.00	1.05	0.99	0.98	0.96
λ_{32}^T	0.08	0.08	0.08	0.08	0.10	0.09	0.08	0.07
ϕ_0^T	0.25	0.54	0.72	0.82	1.02	0.52	0.45	0.33
ν_0^T	2.11	1.50	1.43	1.44	1.46	1.66	1.66	1.69
ν_1^T	–	–	–	–	–	–	–	–
γ_0^T	0.96	0.21	0.02	-0.01	-0.05	0.36	0.39	0.48
λ_0^Δ	0.68	0.58	0.57	0.57	0.54	0.62	0.60	0.63
λ_{11}^Δ	0.87	0.92	0.93	0.93	0.90	0.90	0.92	0.94
λ_{21}^Δ	0.93	0.93	0.93	0.93	0.90	0.93	0.94	0.95
λ_{22}^Δ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
λ_{31}^Δ	1.75	1.75	1.75	1.75	1.77	1.76	1.74	1.72
ϕ_0^Δ	0.04	0.01	0.00	0.00	0.01	0.01	0.01	0.02
ν_0^Δ	1.41	1.32	1.30	1.29	1.32	1.33	1.32	1.34
ν_1^Δ	-0.51	-0.46	-0.44	-0.43	-0.49	-0.47	-0.44	-0.45
λ_0^ρ	0.09	0.09	0.10	0.10	0.09	0.10	0.09	0.08
ϕ_0^ρ	1.50	1.44	1.43	1.42	1.44	1.41	1.46	1.47
ν_0^ρ	0.55	0.55	0.55	0.55	0.56	0.54	0.54	0.57
c_{01}	-0.92	-0.92	-0.91	-0.91	-0.92	-0.81	-0.98	-0.95
c_{02}	0.06	0.06	0.06	0.06	0.09	0.07	0.07	0.01
c_{03}	-0.22	-0.21	-0.20	-0.21	-0.20	-0.20	-0.24	-0.19
c_{11}	1.61	1.60	1.60	1.60	1.56	1.52	1.69	1.64
c_{12}	-0.19	-0.19	-0.19	-0.19	-0.19	-0.20	-0.23	-0.13
c_{13}	0.34	0.33	0.33	0.33	0.27	0.31	0.39	0.33
c_{21}	-1.79	-1.79	-1.79	-1.79	-1.87	-1.84	-1.82	-1.61
c_{22}	0.34	0.33	0.33	0.33	0.39	0.37	0.27	0.29
c_{23}	-0.28	-0.29	-0.28	-0.28	-0.34	-0.31	-0.27	-0.21
c_{31}	1.61	1.60	1.60	1.60	1.56	1.52	1.69	1.64
c_{32}	-0.19	-0.19	-0.19	-0.19	-0.19	-0.20	-0.23	-0.13
c_{33}	0.34	0.33	0.33	0.33	0.27	0.31	0.39	0.33

Table 3.3: Average parameter estimates from fitting the reduced version of model (3), setting $\nu_T^1 = 0$, under the sixteen combinations of choices of smoothing window and block length (four each, indicated by column). Averages are taken across each of the four estimates under the same choice of window length or block length. The parameter estimates are not very sensitive to these choices, but for some parameters there are systematic differences, with the estimate becoming either bigger or smaller as the size of the window or block length increases. These differences are small enough that the effect on the fitted spectra is on average small.

REFERENCES

Amari, S.-i. and H. Nagaoka (2007). *Methods of information geometry*, Volume 191. Providence, RI: American Mathematical Society.

Barnes, E. A. (2013). Revisiting the evidence linking arctic amplification to extreme weather in midlatitudes. *Geophysical Research Letters 40*(17), 4734–4739.

Calder, C. A. (2008). A dynamic process convolution approach to modeling ambient particulate matter concentrations. *Environmetrics 19*(1), 39–48.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics 14*(3), 1171–1179.

Castruccio, S., D. J. McInerney, M. L. Stein, F. Liu Crouch, R. L. Jacob, and E. J. Moyer (2014). Statistical emulation of climate model projections based on precomputed gcm runs. *Journal of Climate 27*(5), 1829–1844.

Castruccio, S. and M. L. Stein (2013). Global space–time models for climate ensembles. *The Annals of Applied Statistics 7*(3), 1593–1611.

Chang, W., M. Haran, R. Olson, K. Keller, et al. (2014). Fast dimension-reduced climate model calibration and the effect of data aggregation. *The Annals of Applied Statistics 8*(2), 649–673.

Chang, X. and M. L. Stein (2014). Wavelet methods in interpolation of high-frequency spatial–temporal pressure. *Spatial Statistics 8*, 52–68.

Collins, W. D., C. M. Bitz, M. L. Blackmon, G. B. Bonan, C. S. Bretherton, J. A. Carton, P. Chang, S. C. Doney, J. J. Hack, T. B. Henderson, et al. (2006). The community climate system model version 3 (ccsm3). *Journal of Climate 19*(11), 2122–2143.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics 25*(1), 1–37.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *The Annals of Statistics 28*, 1762–1794.

Dahlhaus, R. (2012). Locally stationary processes. In T. Subba Rao, S. Subba Rao, and C. R. Rao (Eds.), *Handbook of Statistics, Time Series Analysis: Methods and Applications*, Chapter 13, pp. 351–408. Amsterdam: North-Holland.

Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and M. G. Bosilovich (2012). Evaluation of the reanalysis products from gsfc, ncep, and ecmwf using flux tower observations. *Journal of Climate 25*(6), 1916–1944.

Fanshawe, T. R. and P. J. Diggle (2012). Bivariate geostatistical modelling: a review and an application to spatial variation in radon concentrations. *Environmental and Ecological Statistics 19*(2), 139–160.

Francis, J. A. and S. J. Vavrus (2012). Evidence linking arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters 39*(6), L06801.

Gneiting, T., W. Kleiber, and M. Schlather (2010, September). Matérn Cross-Covariance Functions for Multivariate Random Fields. *Journal of the American Statistical Association 105*(491), 1167–1177.

Gotway, C. A. and L. J. Young (2002). Combining incompatible spatial data. *Journal of the American Statistical Association 97*(458), 632–648.

Goulard, M. and M. Voltz (1992). Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. *Mathematical Geology 24*(3), 269–286.

Guinness, J. and M. L. Stein (2013a). Interpolation of nonstationary high frequency spatial–temporal temperature data. *The Annals of Applied Statistics 7*(3), 1684–1708.

Guinness, J. and M. L. Stein (2013b). Transformation to approximate independence for locally stationary gaussian processes. *Journal of Time Series Analysis 34*(5), 574–590.

Hall, P., J. L. Horowitz, and B.-Y. Jing (1995). On blocking rules for the bootstrap with dependent data. *Biometrika 82*(3), 561–574.

Hawkins, E., T. M. Osborne, C. K. Ho, and A. J. Challinor (2013). Calibration and bias correction of climate projections for crop modelling: An idealised case study over europe. *Agricultural and Forest Meteorology 170*, 19–31.

Ho, C. K., D. B. Stephenson, M. Collins, C. A. T. Ferro, and S. J. Brown (2012). Calibration strategies: A source of additional uncertainty in climate change projections. *Bulletin of the American Meteorological Society 93*(1), 21–26.

Holmes, C. R., T. Woollings, E. Hawkins, and H. de Vries (2015). Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate* (2015).

Ingebrigtsen, R., F. Lindgren, and I. Steinsland (2014). Spatial models with explanatory variables in the dependence structure. *Spatial Statistics 8*, 20–38.

IPCC (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press. Houghton, J. T., and Ding, Y. and Griggs, D. J. and Noguer, M. and van der Linden, P. J. and Dai, X. and Maskell, K. and Johnson, C. A. (eds.).

IPCC (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press. Solomon, S., and Qin, D. and Manning, M. and Chen, and Marquis, M. and Averyt, K. B. and Tignor, M. and Miller, H. L. (eds.).

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press. Stocker, Thomas and Qin, Dahe and Plattner, Gian-Kasper and Tignor, M and Allen, Simon K and Boschung, Judith and Nauels, Alexander and Xia, Yu and Bex, Vincent and Midgley, Pauline M (eds.).

Krafty, R. T., S. Xiong, D. S. Stoffer, D. J. Buysse, and M. Hall (2012). Enveloping spectral surfaces: covariate dependent spectral analysis of categorical time series. *Journal of Time Series Analysis 33*(5), 797–806.

Lawrence, M. G. (2005). The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society 86*(2), 225–233.

Leeds, W. B., E. J. Moyer, and M. L. Stein (2015). Simulation of future climate under changing temporal covariance structures. *Advances in Statistical Climatology Meteorology and Oceanography 1*, 1–14.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The hadcrut4 data set. *Journal of Geophysical Research 117*(D8). D08101.

Neto, J. H. V., A. M. Schmidt, and P. Guttorp (2014). Accounting for spatially varying directional effects in spatial covariance structures. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 63*(1), 103–122.

Neumann, M. H. and R. Von Sachs (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics 25*(1), 38–76.

Ombao, H., J. Raz, R. von Sachs, and W. Guo (2002). The slex model of a non-stationary random process. *Annals of the Institute of Statistical Mathematics 54*(1), 171–200.

Piani, C. and J. O. Haerter (2012). Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters 39*(20), n/a–n/a. L20401.

Poppick, A., D. J. McInerney, E. J. Moyer, and M. L. Stein (2016). Temperatures in transient climates: improved methods for simulations with evolving temporal covariances. *The Annals of Applied Statistics 10*(1), 477–507.

Poppick, A. and M. L. Stein (2014). Using covariates to model dependence in nonstationary, high-frequency meteorological processes. *Environmetrics 25*(5), 293–305.

Priestley, M. (1981). *Spectral Analysis and Time Series: Multivariate Series, Prediction and Control*. New York, NY: Academic Press.

Reich, B. J., J. Eidsvik, M. Guindani, A. J. Nail, and A. M. Schmidt (2011). A class of covariate-dependent spatiotemporal covariance functions for the analysis of daily ozone concentration. *The Annals of Applied Statistics 5*(4), 2425–2447.

Rougier, J., D. M. Sexton, J. M. Murphy, and D. Stainforth (2009). Analyzing the climate sensitivity of the hadsm3 climate model using ensembles from different but related experiments. *Journal of Climate 22*(13), 3540–3557.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, et al. (2010). The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society 91*(8), 1015–1057.

Salazar, E., B. Sansó, A. O. Finley, D. Hammerling, I. Steinsland, X. Wang, and P. Delamater (2011). Comparing and blending regional climate model predictions for the american southwest. *Journal of Agricultural, Biological, and Environmental Statistics 16*(4), 586–605.

Sansó, B. and C. Forest (2009). Statistical calibration of climate system properties. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 58*(4), 485–503.

Sansó, B., C. E. Forest, D. Zantedeschi, et al. (2008). Inferring climate system properties using a computer model. *Bayesian Analysis 3*(1), 1–37.

Santer, B. D., T. M. L. Wigley, M. E. Schlesinger, and J. F. Mitchell (1990). *Developing climate scenarios from equilibrium GCM results*. Max-Planck-Institut für Meteorologie.

Schmidt, A. M., P. Guttorp, and A. O'Hagan (2011). Considering covariates in the covariance structure of spatial processes. *Environmetrics 22*(4), 487–500.

Schneider, T., T. Bischoff, and H. Płotka (2015). Physics of changes in synoptic midlatitude temperature variability. *Journal of Climate 28*(6), 2312–2331.

Screen, J. A. and I. Simmonds (2013). Exploring links between arctic amplification and mid-latitude weather. *Geophysical Research Letters 40*(5), 959–964.

Semenov, M. A. and E. M. Barrow (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change 35*(4), 397–414.

Sham Bhat, K., M. Haran, R. Olson, and K. Keller (2012). Inferring likelihoods and climate system characteristics from climate models and multiple tracers. *Environmetrics 23*(4), 345–362.

Shumway, R. H. and D. S. Stoffer (2011). *Time Series Analysis and Its Applications: With R Examples*. New York, NY: Springer.

Stein, M. L. (2005). Statistical methods for regular monitoring data. *Journal of the Royal Statistical Society - Series B: Statistical Methodology 67*(5), 667–687.

Stein, M. L. (2009). Spatial interpolation of high-frequency monitoring data. *The Annals of Applied Statistics 3*(1), 272–291.

Teutschbein, C. and J. Seibert (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology 456*, 12–29.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research 47*(1), 123.

Ver Hoef, J. M. and R. P. Barry (1998). Constructing and fitting models for cokriging and multivariable spatial prediction. *Journal of Statistical Planning and Inference 69*(2), 275–294.

Vrac, M. and P. Friederichs (2015). Multivariate – intervariable, spatial, and temporal – bias correction. *Journal of Climate 28*(1), 218–237.

Wang, W., P. Xie, S.-H. Yoo, Y. Xue, A. Kumar, and X. Wu (2011). An assessment of the surface climate in the ncep climate forecast system reanalysis. *Climate Dynamics 37*(7-8), 1601–1620.

Wheeler, T. R., P. Q. Craufurd, R. H. Ellis, J. R. Porter, and P. V. Vara Prasad (2000). Temperature variability and the yield of annual crops. *Agriculture, Ecosystems & Environment 82*(1), 159–167.

Wilks, D. S. and R. L. Wilby (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography 23*(3), 329–357.

Williamson, D., M. Goldstein, L. Allison, A. Blaker, P. Challenor, L. Jackson, and K. Yamazaki (2013). History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Climate Dynamics 41*(7-8), 1703–1729.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change 62*(1-3), 189–216.

Yeager, S. G., C. A. Shields, W. G. Large, and J. J. Hack (2006). The low-resolution ccsm3. *Journal of Climate 19*(11), 2545–2566.