Essays on Climatic Extremes, Agriculture and Natural Resource Management

Chandra Kiran B Krishnamurthy

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in Sustainable Development in
the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011

Table of Contents

List of Figures

List of Tables

Dissertation Committee Members

Geoffrey Heal

Finance and Economics Division

Columbia University

Upmanu Lall

Department of Earth and Environmental Engineering &

International Research Institute for Climate & Society

Columbia University

Bernard Salanié

Department of Economics

Columbia University

Wolfram Schlenker

Department of Economics &

Department of International and Public Affairs

Columbia University

Wonghee Tim Huh

Operations and Logistics Division

Sauder School of Business

University of British Columbia

Acknowledgments

I wish to sincerely acknowledge the assistance of my advisors, Geoffrey Heal and Upmanu Lall, on all matters academic over the past four years. Without their support, and especially without Manu's mentorship, advice and practical assistance, this dissertation would have taken far longer than it did. I am indebted to Wolfram Schlenker and Bernard Salanié for productive discussions, advice and encouragement on my research. I must particularly thank Tim Huh for advice on, and encouragement of, my work.

I am especially indebted to my teachers, in particular Eric Verhoogen, Assaf Zeevi, Matias Courdurier, Bhaven Sampat, Sergei Savin, Leigh Linden, Clifford Hurvich (NYU) and William Greene (NYU) for channeling my interest and exposing me to a variety of ways of thought. In particular, I thank Eric Verhoogen and Assaf Zeevi for helpful discussions and honest advice. I also thank the people who have created and sustained the Sustainable Development doctoral program: Mona Khalidi, John Mutter, Joshua Graff Zivin, Matthew Neidel, Jeffrey Sachs, Douglas Almond, Lisa Anderson and Geoffrey Heal; as well as many faculty and researchers who helped me in a variety of ways: Andrew Robertson, Tobias Siegfried, Arthur Greene, Bhaven Sampat, Mark Cane and Balaji Rajagopalan. I am also grateful to Daniel Stellar and Samantha Tress of the Columbia Water Center and to Pann McCuaig for his expert and prompt computer support.

I owe much to my colleagues, in particular, Gordon McCord, Anisa Nwachuku, Marta Vicarelli, Solomon Hsiang, Ram Fishman, Margaret Mcleod, Haitao Ruan, Xiaojia Bao, Nicole Ngo, Kyle Meng and Lily Parshall, for their encouragement, support and motivation. I wish to especially thank Gordon McCord, Ram Fishman and Christina Karamperidou for sharing the joys and frustrations of research, and keeping me motivated about my work and providing me a perspective on its importance.

For the gridded rainfall datasets used in Chapters 2-5, I must thank the Indian Meteorological Department and in particular, Dr Shivananda Pai and Dr Rajeevan Nair. I must also thank Balaji Rajagopalan for his assistance in, and encouragement of, my work

in Chapter 5.

Last, but by no means least, I must thank my parents, brother, and wife, for their constant support and encouragement.

I gratefully acknowledge financial support from Columbia University and the Columbia Water Center. I also thank the Columbia Water Center for providing research facilities.

Dedicated to

My Loving Parents,

To whom I owe a debt too deep for words

And to My Wife, Kalyani,

for her love and support

ज्नानेन तु तदज्नानं येषां नाशितमात्मनः
तेषामादित्यवज्ज्नानं प्रकाशयति तत्परम्

When, however, one is enlightened with the knowledge by which nescience is destroyed, then his knowledge reveals everything, as the sun lights up everything in the daytime.

Bhagvad-Gitā (§5, verse 16)

chapter 1

Introduction

1.1 Motivation

That human societies and systems are impacted by climatic (and geographic) factors are self-evident; indeed, societies have adapted, historically, to their climatological and geographic settings. With increasing technological ability to directly alter, at a large scale, natural systems, post-industrial human societies are increasingly exerting a substantial impact on the environment, thereby establishing an interactive system at a global scale; canonical instances of such impacts include anthropogenic ozone depletion and climate change. Given the self-evident and historical nature of natural impacts on human systems, it is suprising that there appears little rigorous quantification of the impacts of *changes* in such natural systems, either autonomous–such as increasing desertification and changes in rainfall patterns–or anthropogenic, on human societies. Attribution of any specific change in the human system to changes in natural system(s) appear confounded by a variety of factors, including the issue of time scales of relevance.

In particular, two issues appear relevant to understanding the linkages between changes in natural system(s) and their role in causing changes in human systems. First, large natural systems rarely alter at a specific point in time; rather, they undergo a (relatively) slow change in equilibrium, to a newer, possibly less desirable, one. Human societies, over time, perceive such changes and accommodate to them in a variety of ways. Second, changes in the natural systems lead to changes in human systems which are tightly coupled to those natural systems which, in turn, can lead to reorganization of societies, if such human systems are sufficiently important. Thus, attribution of such change in human societies to a particular cause maybe confounded by the time scales and by the internal processes of the human systems themselves.

Anthropogenic climate change appears the most important large-scale (in this case, planetary) change in natural systems in our post-industrial society, and this change is

anticipated to affect a number of aspects of human systems. Yet there is little *quantitative* understanding of the impact of such, possibly momentous, changes upon the human system. Difficulties appear to arise for two reasons; first, there is no one societally meaningful and consistent measure of "changes" in the climate and second, there is no particularly simple representation of the impact of any given change in climate. The latter possibly stems from the fact that the human system is constantly evolving to respond to particular climatic (and other natural) conditions; changes in such conditions, therefore, engender changes in human actions upon which these natural systems impact.

The natural system with the most direct and meaningful impact on human societies, impacted by changes in the climate system is the hydrologic system. This system is expected to undergo substantial alteration, depending upon the magnitude of changes in climatic patterns. Most indications (Trenberth [1999], Bates *et al.* [2008], Huntington [2006] and references therein) point to enhanced precipitation at certain locations and increased aridity at others. Most importantly, increased variability and increasing incidence of extremes of precipitation is clearly foreseen. On the other hand, as regards the human system, there appears substantial agreement (see Smith *et al.* [2007]) that activities directly using the natural system as an input, in particular agriculture, is most likely substantially impacted. In the case of both the natural and the human system, however, there is no clear consensus yet on the magnitude of impacts; indeed, in many cases, even the direction of impacts likely vary regionally. Finally, in both systems, the most important issue is that of "detecting" a signal of change and of understanding the possible impacts of such changes. To a large extent, both are empirical questions, to be investigated using observational data.

Understanding the issue of changing hydrologic extremes may be seen (conceptally at least) as having two components: the first is primarily understanding if, over a reasonably long period of the past, there have been substantial changes in extremes of rainfall; this component provides a basis for the second one, which involves understanding (projecting)

changes in extremes in future climates relative to the historical. The challenge of climate change therefore is two-fold: understanding and summarizing changes in the climate system (and systems driven by it, such as the hydrologic system) and translating such an understanding into quantitative assessments of impacts upon different activities.

The first part of this dissertation attempts to add to our understanding of both aspects (referred to above) of the climate change issue, using tools from two different disciplines: Economics and Hydro-Climatology, for the case of India. Modifications of conventional approaches in Environmental Economics are used to obtain a quantitative understanding of the type of changes in agricultural productivity induced by probable scenarios of climate change. Using a newly available observational rainfall data set from the Meteorological Department, assessments are carried out of changes in extremes of rainfall. These putative changes are then assessed in relation to certain larger scale climatological phenomena which presumably impact hydrologic extremes, based on hypotheses presented in the literature.

There is yet another common thread unifying the first part of the dissertation. Most statistical methods used in the literature on quantifying impacts of climate change on agriculture have focused on mean impacts, implicitly assuming a mean shift of the distribution of yield. Yet, it is not evident that such is necessarily the case; in other words, instead of *testing* whether a hypothesis of climate change leading to a smooth mean impact on yield, holds, it is implicitly presumed that it does. Second, in most studies analyzing observed trends over large regions (such as nation-states or sub-regions), analysis is carried out *after* some manner of aggregating rainfall (such as in Goswami *et al.* [2006]) is chosen. In both cases, interest centers on presumably *pooling information* at various locations (district, in the former, and grid, in the latter) in order to make inference regarding the larger region.

In the case of climate change impacts on yields, consider the following two scenarios regarding the distribution of yield[1]: in the first, the upper and lower quantiles change (say

[1] Assuming that the quantiles are fixed over time.

increase and decrease respectively) with a very minor change in the mean while in the second, the upper quantiles increase substantially while the lower decreases, with the mean impact being the same as in the first scenario. Clearly, both scenarios have very different implications for policy and welfare, while they appear identical from the mean regression perspective, which presumes parallel shift of the quantiles.

Consider now the case of *pooling* rainfall data from a number of grids (stations) by applying a uniform threshold (as in Goswami *et al.* [2006]; Rajeevan *et al.* [2008]). If grids are sufficiently heterogeneous, then not only are different grids likely to be selected for different years but (as in the case of the data we use) only a few grids contribute to the analysis. Thus, this type of pooling leads to only the grids with sufficiently high *average* level of rainfall contributing to the analysis.

In both cases, the analysis carried out here *allows explicitly* for heterogeneity in the individual unit of observation, when inferring change; in the former, it is heterogeneity of impact on yield (of a given change in climate) while in the latter, it is accounting for heterogeneity of level of rainfall[2] when defining thresholds from which an extreme event is defined. In the case of climate change impact on agriculture, a given change in climate/weather (temperature and rainfall, for instance) is allowed to impact yields at different points on the distribution of yield differently[3]. In the analyses on evaluation of trends over India, attempts are made to account for heterogeneity by applying a different threshold for defining an extreme; this way, the spatial heterogeneity in rainfall levels are accounted for.

India, as a large developing nation, has an agricultural sector which is large, consituting up to *18%* of GDP and *54%* of employment, with a large proportion of subsistence farming.

[2] In other words, extremes must be measured relative to a grid's (station's) own level of rainfall, rather than the rainfall level of some pooled "grid", which is not representative of any given grid.

[3] Put another way, a given increase in temperature is allowed to impact yields very differently, including difference in signs, at different "regions" of the distribution of yield. To be more concrete, an increase in temperature is allowed to be beneficial at the upper (lower) quantiles and detrimental at the lower (upper) quantiles. Whether they are indeed similiar is indeed an empricial question and is explicitly tested for.

It is among the largest food producers in the world (second and third-largest producer of rice and wheat respectively), and yet suffers from both, a high rate of lack of access to stable food supply and fluctuations in the production of food driven by variations in large scale monsoonal features (reflecting its largely rainfed and subsistence farming). These features, compounded by a lack of effective public food distribution systems, renders the poorer population vulnerable to fluctuations in production of food grains.

It has a hydrological cycle which is complex, driven primarily by the monsoon phenomenon, with substantial inter- and intra-year variability in rainfall patterns. Thus, while average all-India rainfall is exceedingly high (at 1160 mm; see Kumar *et al.* [2005]), the spatial distribution of rainfall and its utilizability is such that India has among the largest semi-arid regions in the world, particularly in the South-Central and North-Western regions. The seasonally intermittent (about *80%* of annual rainfall occurs in less than *100* days) and intense rainfall (in many cases, more than half the annual rainfall occurs in less than *100* hours) during the monsoon season poses hazards of both periodic drought and flooding. In particular, rainfall and tropical-cylone related extreme events are a significant hazard in much of the Indian sub-continent (encompassing India, Bangladesh, Nepal and Pakistan). Rainfall extremes, therefore, have significant and large-scale societal impacts through agricultural production, flood hazards etc.

Thus far, we have addressed the issue of impacts of changes in natural environment upon the human system. There is however, its obverse, human management of the natural system. Human societies use and interact with, directly or indirectly, a very large proportion of the natural system. In particular, the economic system uses, as its input (or as a sink, as in the case of pollutants, including green house gases) natural resources, including minerals, water and the atmosphere. Given the importance of such natural inputs in human production, these have been substantially studied and modelled. In particular, there is a large Economics literature on “optimal” (from the persepctive of the economic system) management of a

variety of natural resources. Yet, human management of resources as diverse as groundwater aquifers and fisheries has been far from optimal even from its own perspective, and has led to substantial overuse of these systems (see Wada *et al.* [2010] and references therein for groundwater depletion, Worm *et al.* [2009] for the case of fisheries). In particular, groundwate aquifers throughout a large part of the semi-arid regions around the globe have been subject to overdraft.

Recharge of aquifers occur primarily through rainfall or through irrigation and riverine flows; when extraction is larger than recharge, as is true in a variety of semi-arid regions around the world, there is an effective mining of water. Rapid depletion of these resources thus has an increasing impact through a positive feedback: depletion raises the cost of further extraction, rendering such extraction, afer a point, *uneconomical* societally (see Dasgupta [1982] for a detailed discussion). While this phenomenon is likely to prevent *physical* exhaustion/extinction of a resource, the impact on production systems dependent on such inputs (such as agricultural systems) are likely eventually to be substantial, especially in the developing nation setting. For instance, food prices are likely to rise[4], adversely impacting large populations of poor; in addition, it may also further increase the ranks of the poor by depriving large numbers of small farmers their means of livelihood. Thus, there are two aspects to the issue: first, understanding the specific context of managing groundwater in the semi-arid, developing nation context, and second, exploring creative and context-specific solutions to the issue.

In this context, the second part of the dissertation addresses the issue of unified, coherent and more realistic analysis of groundwater-based models in the Economics literature. The attempt here may be viewed as a modest attempt at addressing the first of the two aspects.

[4] A large part of the gains in food production in monsoonal semi-arid nations such as India, China and Bangladesh are directly as a result of explosion in groundwater-based agriculture.

1.2 Structure and Scope

The dissertation comprises of five individual papers, which span the two dimensions indicated in Section 1.1. The methods used and the questions addressed are, by design, diverse. To summarize, three distinct, yet related, issues are addressed:

1. Quantification of climate change impact on agriculture
2. Assesment, and attribution, of trends in extremes of rainfall
3. Structure of groundwater management models

We indicate below the nature of the issues addressed, the methodologies used and the inter-relationships between them. Chapter 2 addresses the question of probable impact of climate change on Indian agriculture. The question here relates to an understanding of a direct impact between changes in the natural system and a human system directly dependent upon it. This may be better understood conceptually in the following framework; let $Y_{t,i}$ denote the yield in district i in year t of a given crop, $\{X_{i,t,k}\}$ denote a "climate" field or alternately, a set of "climate" variables. Nonetheless, denoting by $Q_{Y_{i,t}}(\tau|X_{i,t,k})$ as the quantile function of the (random) variable $Y_{i,t}$, where $\tau \in (0,1)$ is a quantile of the distribution of $Y_{i,t}$, we $Q_{Y_{i,t}}(\tau|X_{i,t,k})$ as being linear i.e.

$$Q_{Y_{i,t}}(\tau|X_{i,t,k}) = \beta(\tau)X_{it}$$

where $\beta(\tau) = (\beta_1(\tau), \ldots, \beta_k(\tau))$. The case of the mean regression corresponds to $\beta(\tau) = \beta,\ \forall\tau$.

While partly based on existing methdologies in Econometrics, the work here represents a departure from the existing work in both Environmental Economics and agricultural modeling in a variety of dimensions. While empirical and using a reduced form approach to data analysis, the methodology adopted nonetheless is flexible enough to allow both, a

more detailed characterization of impacts as well as accounting of impacts of sub-seasonal variability in monsoon rainfall.

To understand Chapters 3-5, consider the set up of a statistical inverse problem. We emphasize that the explanation below is intended to be an informal elucidation of ideas, since we do not use the formal tools of this theory. Let Θ index the class of models, with Θ some space (see Evans and Stark [2002] for a full and formal development of the set up), with the model with index$\theta \in \Theta$ being referred to as the "model θ". The *forward mapping* is then the mapping $\theta \mapsto \mathbb{P}_\theta$, with $\mathbb{P}_\theta$ the probability distribution for the observables. The inverse problem refers to a scenario wherein data X drawn from $\mathbb{P}_\theta$ are oberved, for *unknown* θ; we then wish to use the knowledge of X and the fact that $\theta \in \Theta$ to estimate $g(\theta)$ where g is a function defined on Θ.

In our case, let t, s denote respectively time and space (grid location) and let $\tilde{\theta}$ denote our "signal" over space and time, where the *signal* is a pattern or a trend over a region across a given time period. Given $\{X_{t,s}\}_{t\in T, s\in S}$, our objective is to recover the physically relevant "signal" $\tilde{\theta}$. In other words, we wish to estimate $\tilde{\theta}$ from the relationship

$$\{X_{t,s}\} = g(\tilde{\theta})$$

In Chapter 3, the relationship of interest is existence of a trend over a large spatial domain. In other words, interest centers on understanding whether, over a large domain (all of India, in fact) there are any discernible increasing (or decreasing) trends in rainfall extremes. In order to infer this, we use gridded rainfall data for 53 years, compute a non-parametric estimate of a (deterministic) trend and then make (statistical) inferences about whether a trend exists over all of India, as conjectured in the literature.

In Chapter 4, the relevant signal of interest is again a trend, now in only one metric (frequency of exceedance of a certain threshold) over a smaller region of India. The differences from the problem in Chapter 3 may again be illustrated in the notation of the

statistical inverse problem setting above. Denote by $\{Y_t^k\}_{k\in K}$ data on a set of K climatic predictors or precursors, which potentially impacts rainfall in our region of interest. The relevant inverse problem now is:

$$\left\{X_{t,s}, Y_t^k\right\} = G_k\left(\tilde{\theta}\right)$$

In other words, the signal extraction problem is more difficult, in that there are additional "predictors", climatic phenomena which can mask or otherwise contaminate the signal $\tilde{\theta}$. In view of the added difficulties, the major difference from Chapter 3 is in the methodology used. Explicitly recognizing the discrete (count) nature of the data $\{X_{t,s}\}$, and to account for predictors, we use a new regression setting which, in addition, filters only areally large signals, quantile regression for count data.

Finally, Chapter 5 reports on an ongoing investigation, to understand which consider now the following inverse problem:

$$\{X_{t,s,\omega}\} = g(\widetilde{\theta_\omega})$$

where $\omega \in \Omega$ is some frequency (as in the physical frequency or inverse of the time period of a relevant physical cycle) of interest, in the frequency spectrum Ω. In other words, the inverse problem now corresponds to recovering a pattern of behavior of extremes which now *varies by frequency.* The problem dealt with in Chapter 1 is a sub-set of this problem and may be obtained by setting $\omega = 0$, where a deterministic trend is seen as a spike in the frequency spectrum at 0 frequency. We investigate four frequencies of interest, the *ENSO* (El Niño Southern Oscillation), *QBO* (the Quasi-Biennial Oscillation) and *decadal* frequencies, in addition to the trend, at $\omega = 0.3,\ 0.4,$ and 0.1 respectively.

The final substantive chapter, Chapter 6, deals with understanding the structural properties of groundwater models. To assist in clearly visualizing the objectives of the analysis,

as well as the contribution of this chapter to the existing literature, we begin with some notation. Consider first a simple, single-cell aquifer, possibly a "bath-tub" type aquifer. Let x_t denote the volume of water in the aquifer at the beginning of time period $t \in \{1, 2, \ldots, T\}$, where time is discrete, and for now $T < \infty$ and w_t denote the extraction (pumping) of groundwater in period t. The farmer observes x_t, makes a decision about w_t and then random recharge (in the form of rainfall) occurs. Thus, the next period stock is, simply, $\tilde{X}(x_t, w_t) = x_{t+1} = x_t - w_t + R_t$. Denote by $\Pi(x_t, w_t) = B(w_t) - C(x_t, w_t)$ the net benefit to the farmer of pumping in period t, where $B(w)$ are concave and increasing benefits, with $C(x, w)$ being the cost of pumping. Following the literature (cited below), and in keeping with hydrologic facts, we take that $C(x, w)$ is lower for higher levels of stock (for any level of pumping), and increases in pumping (for any given stock).

The decision framework just set up is a dynamic framework, with short-term benefits of increased extraction offset by longer-term benefits of reduced costs. This decision problem is expressed, in this chapter, as a formal discrete-time Stochastic dynamic programming problem. In related literature in Environmental Economics (Burness and Brill [2001]; Burt [1964, 1966]; Gisser and Sanchez [1980]; Knapp and Olson [1995]; Provencher and Burt [1993]; Zeitouni [2004], among many others), the important questions of interest are: Is it necessarily true that

1. $w_t(x)$ is increasing in x_t?
2. $w_t(x)$ is decreasing in t?
3. $x_{t+1} = x_t - w_t(x)$ is increasing in x_t?

Each of these are intuitive properties, and if they do hold, help in intuition. Typically, smoothness assumptions (*joint* concavity of Π) are imposed on the problem in order to yield these desirable properties. Yet, in many settings for $C(x, w)$, including the most commonly used ones, these *do not hold*. In other words, these properties, in the framework

just described, have not been shown to hold for even the most simplifed form of $C(x, w)$ used in the literature. In this setting, we generalize the cost function in two dimensions, and for each of these three forms of $C(x, w)$, prove that some of the three properties listed (and two which are not) hold. We do not impose any smoothness assumptions on the primitives in this chapter and work in a lattice-theoretic set up.

Second, we also prove that many of these properties survive when risk aversion of the decision maker is taken into consideration ie. when the objective is to maximize $U(\Pi(x, w))$, for the log utility function. We note that, apart from Knapp and Olson [1996] (who work in a different set up) there is no mention of this property in the literature. In addition, we also investigate limiting (in t) behavior of the problem, and show that certain important properties do survive this limiting operation.

Finally, we investigate the issue of when the Markov Chain $x_{t+1} = x_t - w_t + R_t$ converges to a unique stochastic steady state[5] when property (3) does not hold. We use a very powerful, and little used (in the Environmental Economics literature), result to illustrate that, for a very wide class of models, there does exist a unique invariant distribution. Our work in this Chapter thus unifies, generalizes and proves a variety of properties for a large class of stochastic models used in this literature.

[5] Analogous to the converence of a deterministic dynamical system to a deterministic steady state, a stochastic dynamical system, or equivalently, the Markov Chain, in (3), can converge to a steady state probability distribution for X_t.

Bibliography

B. Bates, Z.W. Kundzewicz, S. Wu, and J. Palutikof. Climate change and water. *Climate change and water*, 2008.

H.S. Burness and T.C. Brill. The role for policy in common pool groundwater use. *Resource and energy economics*, 23(1):19–40, 2001.

O. R. Burt. Optimal resource use over time with an application to ground water. *Management Science*, 11(1):80–93, 1964.

O. R. Burt. Economic control of groundwater reserves. *Journal of Farm Economics*, 48(3):632–647, 1966.

P. Dasgupta. *The control of resources*. Harvard Univ Pr, 1982.

S.N. Evans and P.B. Stark. Inverse problems as statistics. *Inverse Problems*, 18:R55, 2002.

M. Gisser and D. A. Sanchez. Competition versus optimal control in groundwater pumping. *Water Resources Research*, 16(4):638–642, 1980.

B.N. Goswami, V. Venugopal, D. Sengupta, M.S. Madhusoodanan, and Prince K. Xavier. Increasing trends of extreme rain events over india in a warming environment. *Science*, 314:1442–1444, December 2006.

T.G. Huntington. Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology*, 319(1-4):83–95, 2006.

K. C. Knapp and L. J. Olson. The economics of conjunctive groundwater management with stochastic surface supplies. *Journal of Environmental Economics and Management*, 28(3):340–356, 1995.

K.C. Knapp and L.J. Olson. Dynamic resource management: Intertemporal substitution and risk aversion. *American Journal of Agricultural Economics*, 78(4):1004, 1996.

R. Kumar, R.D. Singh, and K.D. Sharma. Water resources of India. *Current Science*, 89(5):794–811, 2005.

B. Provencher and O. Burt. The externalities associated with the common property exploitation of groundwater. *Journal of Environmental Economics and Management*, 24(2):139–158, 1993.

M. Rajeevan, J. Bhate, and AK Jaswal. Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. *Geophys. Res. Lett*, 35, 2008.

P. Smith, D. Martino, Z. Cai, D. Gwary, H. Janzen, P. Kumar, B. McCarl, S. Ogle, F. O Mara, C. Rice, et al. Agriculture. Climate change 2007: mitigation. Contribution of working group III to the fourth assessment report of the intergovernmental panel on climate change. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, 2007.

K.E. Trenberth. Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change*, 42:327–339, 1999.

Y. Wada, L.P.H. van Beek, C.M. van Kempen, J.W.T.M. Reckman, S. Vasak, and M.F.P. Bierkens. Global depletion of groundwater resources. *water resources*, 37(L20402), 2010.

B. Worm, R. Hilborn, J.K. Baum, T.A. Branch, J.S. Collie, C. Costello, M.J. Fogarty, E.A. Fulton, J.A. Hutchings, S. Jennings, et al. Rebuilding global fisheries. *science*, 325(5940):578, 2009.

N. Zeitouni. Optimal extraction from a renewable groundwater aquifer with stochastic recharge. *Water Resources Research*, 40(6):W06S19, 2004.

chapter 2

The Distributional Implications of Climate Change on Indian Agriculture: A Quantile Regression Approach

2.1 Introduction

Given the incontrovertible evidence on human-induced changes in climate, research is increasingly focussed on better and finer-scale understanding of the pathways of the impacts of such changes on economic activities. Quantifying the nature of such impacts, under a variety of possible scenarios of climate change, is a critical stage in the process of understanding the nature and magnitude of adaptations to the changes. The impacts of climate change are likely evident on a variety of time scales and climatic variables, and research to date has focused mostly on the manifestations of such change, primarily through changes in precipitation and temperature. Many regions of the globe are anticipated to undergo rapid warming at the surface, with a possibility of increasing precipitation.

Of the many sectors of an economy impacted by climate change, agriculture and forestry have been identified to be most susceptible, given that weather is a direct component of the production in these sectors. Another reason for the importance of agriculture is its disproportionate share in employment in developing nations[1]. Most studies of climate change impacts on agriculture focus on the developed world, primarily the US (Deschenes and Greenstone [2007]; Schlenker and Roberts [2006, 2009]; Schlenker *et al.* [2006]), but the importance of the agricultural sector in these regions are smaller than those in the developing nations. Estimates of the impacts of climate change on these economies will presumably be a first and important step in adapting to these changes. The importance of the task is magnified by the predictions of both larger changes in climate (and therefore weather) and larger uncertainties in those changes in the developing nations (predominantly in the tropical and sub-tropical regions) than in the more developed (and temperate) ones. This paper provides direct evidence on the impact of climate change on Indian agriculture, and finds that climate change is likely to reduce yields of many food grains substantially,

[1]The share of agriculture in employment, for India, is about 54%.

although there is also intriguing evidence of only small negative and even slightly positive changes for rice.

Previous studies quantifying the impact of climate change on agriculture fall into one of three categories, based on the approach used to quantify such impacts: production function approach, the Ricardian approach and a panel data approach. The production function approach is based on controlled agricultural experiments in a laboratory-like setting, under agronomically optimal conditions. As a result, they reflect optimal outcomes, with only minimal adaptation to changes possible, and likely differ from realized outcomes at the farm level. The Ricardian approach, pioneered by Mendelsohn *et al.* [1994], attempts to control for the full range of farmer adaptations possible, by estimating a functional relationship between land values and other (appropriate spatial) climate variables (along with other controls). The main premise of the approach is that, with well functioning land markets, land prices will reflect the present discounted value of profits from all uses of land, controlling for many potentially unobservable characteristics. Naturally, the reliability of such methods depend on the tenability of such assumptions (for evidence on this approach to agriculture in the US, see Deschenes and Greenstone [2007]; Fisher *et al.* [2009]).

A variant of this approach, which is less demanding in terms of well functioning markets, is the semi-Ricardian approach, developed in Sanghi *et al.* [1998] and applied to the cases of Brazil and India. This approach uses average profits (over a long period of time) at the district level as a proxy for land prices, with the assumption that land prices, if available, would be the present discounted value of such profits. Using this approach, Sanghi *et al.* [1998] and Kumar and Parikh [2001] find significant negative impacts on profits under a variety of climate change scenarios for India.

The panel data approach uses presumably random year-to-year variation in weather across US counties to estimate the impact of weather on agricultural yield (Schlenker and Roberts [2006, 2009]; Schlenker *et al.* [2006]) and profits (Deschenes and Greenstone [2007]).

This approach overcomes the omitted variable-type issues encountered in the Ricardian approach, and permits controlling for unobserved individual heterogeneity (via the use of county fixed effects), such as farmer/soil quality, while at the same time allowing for some farmer adaptation, providing presumably better estimates than the production function approach. A major drawback with this approach is an inability to reflect changes in technology and choices (such as switching of crops, exit from agriculture, change in cropping seasons etc). Using this approach for India, along with gridded reanalysis weather data, Guiteras [2008] explores the impact of climate change on annual profits at the district level, and estimates moderate losses (up to 9%) for the medium term, and significant (up to 25%) losses for the long term, under a variety of climate scenarios.

It is also important to notice that all of the preceding approaches model the *mean change* in the underlying variable (yield, land value, average profit) by using a mean regression framework (parametric or nonparametric), in which a major assumption implicit is that covariates (weather, climate, controls, etc) impact only the mean of the underlying variable, and do not alter the shape (or scale etc) of the conditional distribution of yield.

In other words, implicit in such an approach is the idea of climate change leading to mean shift in agricultural outcomes, with no changes of underlying relationship between agricultural outcomes and climate. This is problematic for several reasons; for instance, in much of the scientific literature on climate change, focus is on changes in variability (especially in the hydrologic cycle, which determines both long and short-run availability of water supply, a critical ingredient in agriculture) as a result of an altered climate; while such changes are encapsulated in a "mean effect" framework, they are not restricted to it. Research focus in the climate literature therefore is on changes in extremes and in variability changes. Further, the assumption of a mean shift is arguably a *testable hypothesis* in a statistical framework, rather than an implicit axiom.

This paper develops a novel panel data approach for linking observed weather outcomes

and agricultural outcomes, at the district level[2], at any number of chosen quantiles. In this framework, yield, at each quantile, is modeled as being a (potentially nonlinear in covariates) function of weather variables, primarily Growing Degree Days and seasonal (and monthly) rainfall and district-specific fixed effects. The use of district-specific fixed effects allows unobserved determinants of district-specific, time-invariant effects (such as soil type, topography etc) to be controlled for. The estimation strategy is to regress yield on observed weather outcomes and district-fixed effects, for each quantile of interest. Extensions of the framework to flexible modeling of covariate impacts, as well as to fully non-parametric settings using the penalized splines approach in Koenker *et al.* [1994] are also indicated. The framework developed here may therefore be seen as a generalization of the non-parametric estimation outlined in Schlenker *et al.* [2006] to an estimation of various features of the conditional distribution, apart from the mean.

Given that the quantile function completely characterizes the distribution function of a random variable, the idea underlying this approach (and quantile regression in general) is to parametrically obtain estimates of local features of the full conditional (upon covariates) distribution of yield, without sacrificing the interpretability and computational simplicity of the regression framework[3]. Unlike in the linear panel data case, this approach allows for two distinct types of (unobserved) heterogeneity: district-specific unobserved heterogeneity (due to the inclusion of the fixed effects), and heterogeneity of covariate effects (due to the quantile nature of the regression). Region-specific cubic time trends are also included in order to disentangle the impacts of a potentially warming climate over the latter period of the twentieth century and improvements in agricultural productivity over the same period.

[2]The district is the smallest administrative unit in India, and is analogous to an American county, although generally larger.

[3]For instance, alternative and non-parametric estimation of conditional quantile functions are detailed in Koenker [2005, Chapter 7]. These frameworks however are (a) not easily extended to higher dimensions (of covariates) (b) lack interpretability, especially in multi-variate settings (c) not readily generalized to fixed effects settings and (d) computationally burdensome, especially since inference is bootstrap-based.

The predicted impact of climate change at each quantile is computed as the differences in predicted yield at the historical weather and weather realizations from climate models for various climate scenarios (or uniform changes in temperature and rainfall applied to current weather).

This paper uses a unique and recently available dataset of gridded temperature and rainfall data for India, from the Indian Meteorological Department, spanning 1969-1999, and an agricultural dataset, spanning 1971-2005, for agricultural yield at the district level, for estimation. Rice and wheat are grown in different seasons, with very different characteristics, and projections of climate changes are very different for these two seasons. It is therefore important to account for considerations of separate growing seasons and therefore very different impacts on crops in the different growing seasons[4]. This paper find significant negative impacts, of up to 6% of yield, for wheat, for very reasonable scenarios (up to 0.5^0C increase in temperature), and up to *10%* losses for larger increases (1^0C increase in winter temperature).

For rice, however, results indicate minimal losses at all but the highest quantiles considered, and even mildly positive impact in many regions, primarily the Eastern region, a result which is consistent with those in Guiteras [2008], and with a recent study of Welch *et al.* [2010][5]. These results are robust to both flexible functional forms (quadratic, orthogonal polynomials and fully non-parametric) and to different plausible thresholds, for Growing Degree Days.

[4]There are two major growing seasons in India; the summer monsoon growing season, called *kharif*, and the winter growing season, called *rabi*. There are actually two monsoon seasons over India, the most important being the summer or South Western monsoon season, from June to September, during which about *70%* of annual rainfall occurs. The second monsoon season (North Eastern monsoon) affects primarily South-Eastern and North-Eastern regions of India, of which agriculture in South-Eastern states of India (Andhra Pradesh and Tamil Nadu) is significantly impacted. In this paper, we will use terms "summer growing season" ("winter growing season") and "kharif season" ("rabi season") interchangeably.

[5]This study reported an increase in rice yield across different locations under increasing maximum temperature, although at the only study location in India (in Tamil Nadu, South-Eastern India) temperatures reported rarely exceeded 34^0C.

Prior approaches to estimating impacts of climate change for India suffer from two drawbacks: first, given that there are two agricultural seasons, it is not clear that ignoring completely weather(climate) in this season while including agricultural outcomes for this season as part of the response (as in Sanghi *et al.* [1998], Kumar and Parikh [2001] and Guiteras [2008]) provides sensible estimates. This issue aside, it is not clear how to interpret the monetary equivalent of crop yield when summed over crops grown in different growing seasons (i.e. $\sum_k y_{itk}p_{ik}$ where k denotes the crop, y_{itk} the yield of crop k in year t and district i, prices p_{ik} are *fixed* in time) at the district-level (Guiteras [2008]). For any interpretation of this variable as potential productivity, one must also control for all inputs, which is not possible in this case (due to the absence of such data). A final issue is the possible confounding of price impact from productivity i.e. results are likely sensitive to impacts of the specific price used[6]. The approach outlined in this paper avoids the issues raised above altogether by working directly with changes in physical crop yield, addressing crops grown in different seasons separately, and accounting for substantial changes in seasonal weather directly.

The plan of the paper is as follows: we first detail the agronomic and economic basis of the model, turn then to outlining the advantages of quantile regression in the current setting, indicating the benefits and interpretation of fixed-effects in quantile regressions, proceed to outline the econometric strategy adopted and finally, report results of the estimation, including impacts of climate change on yields.

[6]It is also not evident that fixing prices at the *1960* level provides any remedy to the problem outlined for the reason that the year chosen(1960) pre-dates all significant agricultural policy changes in India, at both the national and the state level (such as support prices for rice and wheat and sugar co-operatives for sugarcane, changes in electricity pricing for irrigation and water pricing for surface irrigation). These prices therefore are unlikely to be relevant for later periods.

2.2 Model

The objectives of the paper are to be able to obtain estimates of the impact of weather on crop yields, using a newly available high quality gridded weather data set, and using a newly developed framework for quantile regression for panel data with fixed effects. The approach outlined here allows estimation of the potentially non-linear relationship between temperature and rainfall on crop yields, at different quantiles of yield.

2.2.1 Temperature and Yields

Given the reduced form nature of the statistical approach proposed here, it is imperative to obtain a seasonally aggregate metric of temperature relevant for crop growth (given that aggregate metrics such as "average temperature" yield almost no relevant information). Following the approach outlined in Guiteras [2008]; Schlenker and Roberts [2009], we postulate a temperature limit within which heat is beneficial to plants, in a linear and time-separable manner. The upper and lower limits of temperature are a matter of some debate even in the context of US agriculture (for which they were first derived) and are generally unknown for the Indian context. However, following Guiteras [2008], we fix these limits to be $8°C$ and $34°C$ for the summer growing season. Following Schlenker *et al.* [2006], who illustrate the superiority of the agronomic measure of Growing Degree Days (GDD), we compute this measure using two different methods. In the first case, which we term "Linear Degree Days", the computation is straightforward[7]. Given daily average temperature, T, compute

[7]Degree days were also computed using an alternative method of Baskerville and Emin, with very similar results (not reported here).

$$D(T) = \begin{cases} 0 & \text{T}\leq T_{low} \\ T - T_{low} & T_{low}\leq T \leq T_{upper} \\ (T_{upper} - T_{low}) & T \geq T_{upper} \end{cases} \tag{2.1}$$

where the upper and lower thresholds are set to $8°C$ and 34°C for the summer growing season.

The Linear GDD metric was computed for each day, summed over the growing season (*kharif* and *rabi* seasons) to obtain a seasonal sum of GDD[8]. The fixing of the growing season is less to do with the endogeneity of farmers' sowing and harvesting decisions (which are, for the rainfed areas, dependent on expected onset of the summer monsoon) and more to do with bio-physical conditions, including availability of water. In order to account for potentially harmful effects of temperatures beyond $34°C$ upon plant growth, we also compute "Harmful Degree Days" as the difference between $34°C$ and the daily average temperature.

The choice of temperature limits for the *rabi* season was made more difficult by having little India-specific literature on the topic (much of the literature use only a lower bound and do not use a specific upper bound). Given however the significant temperature constraints on wheat, it was decided to choose $6°C$ as the lower threshold (temperatures only very rarely go much below this over much of the wheat growing regions) and an upper threshold of $28°C$. Choosing slightly different upper thresholds ($27°C$ and $29°C$) did not appreciably alter the results. Harmful Degree Days were computed as in the case of the *kharif* season.

Two points regrading this formulation are worth emphasizing : (a) linearity of plant growth as a function of temperature during the growing season (b) the complete indepen-

[8]In other words, $GDD_k = \sum_{t=1}^{T_k} D(t)$ i.e. the growing season degree days are the sum of the daily degree days over the growing season, for a given year; season length is 122 days for the summer growing season and 183 days for the winter season.

dence between temperature and precipitation, when used in the regression framework with yield (not its transformations) as the dependent variable. For instance, in the regressions in Guiteras [2008] (see for instance, Equation (5), pp 22), this implies the complete separation of the effects of temperature and precipitation upon yield, which is somewhat unrealistic in the case of the summer growing season.

2.2.2 Precipitation and yield

The focus in most of the econometric research on estimation of climate change impacts has been on estimation of temperature impacts upon plant growth. Agronomic and climate literature for India however, has tended to focus, for rainfed agriculture, on the significance of summer monsoon characteristics. However, in the case of India, even without climate change, climatic variability, especially the variability in monsoon rainfall, has been of major interest for agriculture. This is due to (a) the season: the earlier part of the summer growing season from May-Sept are also the hottest months of the year and temperatures generally reduce only with rainfall (b) lack of water availability: seasonal characteristics and the largely rainfed nature of agriculture in India (current estimates indicate that only about *40-50*% of agriculture in India is fully irrigated) means that sowing operations are critically dependent on the arrival of rainfall.

The significance of monsoon rainfall for Indian agriculture is well documented in both, Agricultural and Development Economics (for instance, Rosenzweig and Binswanger [1993]) and in the Agriculture and Climate literature (Gadgil, 2000 etc). In these regions, it is well known that the major constraint is availability of water during the critical sowing period, May to June, and therefore, onset of the monsoon is of some importance. Further, contrary to popular perception of the monsoon as marching South-West to North-West, the monsoon has what are known as “break periods”, periods during which there are long dry spells (of up to a fortnight or more in duration) over large parts of India. There is some speculation

on the harmful effects of these periods, but little actual evidence about its importance for agriculture.

If (as in Guiteras [2008]) identification is based on weather variations from district mean weather, we argue that it is important to (a) account for rainfall variations which are not captured by monthly totals (b) condition monthly and seasonal total rainfall on onset and dry spells. The latter accounts for the fact that years with late onset and years with normal rainfall are not exclusive and therefore, two years with very similar total rainfall can have very different onset days; further, given the concentrated nature of monsoonal rainfall in India, years with a large number of "dry spells" can easily be years with average monthly rainfall (especially true in the semi-arid parts of India with highly variable rainfall patterns). In other words, monthly rainfall is not a sufficiently fine proxy for rainfall distribution within the month, given the relatively concentrated nature of rainfall in India. However, total seasonal rainfall may be a somewhat inadequate measure of such variability due to the concentrated nature of rainfall in many regions in India (for instance, more than 80% of total rainfall in India is documented as occurring in less than 100 hrs). We therefore use more direct measures on the two aspects of monsoon most relevant for agriculture: onset of the monsoon and the number of dry spells during the season. Monsoon onset is defined simply as the first day after April 1 that rainfall on 5 successive days is above 5mm without being followed by 5 successive days with rainfall below 5 mm. Defining as a "dry spell" any 7-day period with rainfall below 5mm, we obtain a count of such spells for every district for every year (see Appendix A for more details on these variables).

To summarize, we use the following variables for summer monsoon season measures of rainfall: seasonal (monthly) total rainfall and onset day, dry spells. For the winter season we use, in addition to seasonal rainfall, a bi-monthly rainfall sum.

2.3 Econometric Strategy

We turn now to characterizing our econometric strategy. We begin with a brief interpretation of linear quantile regression as an extension of linear regression to the heterogeneous effects of covariates case, and then turn to indicating the major ideas that underlie an extension of quantile regression to the fixed effects setting.

2.3.1 Quantile Regression and heterogeneous covariate effects

Consider the so-called location-scale model:

$$Y_{it} = X_{it}\beta + (X_{it}\gamma)\sigma_{it} \tag{2.2}$$

where where σ_{it} is an error term (not necessarily i.i.d) and $\gamma \neq 0$ represents "heterogeneity" (observed). It is evident that the condition $\gamma = 0$ is sufficient for the conditional quantiles to be parallel to one another, and this is generally a *maintained assumption* in mean regressions, parametric or non-parametric. In other words, most mean regression frameworks *assume* that the conditional quantiles are all parallel to one another and that the covariates exert an effect only on the mean of the distribution. Evidently, this assumption is restrictive. In this context, quantile regression maybe viewed as an extension of regression to the "location-scale" models i.e. covariates also affect the scale parameter (variance, for the normal distribution) of the conditional distribution of the response, due to which the quantiles are no longer parallel to one another[9].

For instance, in the current application, it is quite plausible that weather covariates have differential impacts on distinct quantiles of the yield[10]; further, if the effect of climate

[9] In the case of the simple location-scale model above, note that the conditional variance of Y is now a function of X and of γ.

[10] A very common economic application of quantile regression is in Labor Economics, in the analysis of differential impacts of covariates on different parts of the wage distribution. See Fitzenberger *et al.* [2002]

change for India are likely manifested in increases in variance along with changes in mean (as is currently speculated, especially for precipitation), then impacts computed based on mean regression are likely to understate the true magnitude. Similarly, if there is a change in seasonal distribution of rainfall, computing the mean impact will tend to smooth out variations in yield at different quantiles.

2.3.2 Quantile regression for fixed effects panel data models

There is an extensive literature in Economics, particularly in fields such as Labor Economics, on the usage of quantile regressions (see for instance Fitzenberger *et al.* [2002]). Effectively, linear quantile regression is the simplest parametric framework for a scenario in which covariates are expected to exert differential impacts on the distribution of the response.

Panel data with fixed effects allow applied researchers to base identification purely on "within" variation in the covariates. Further, this approach allows arbitrary correlation between the covariates and the fixed effects. There has been, however, little research on the intersection of the two methods. Two approaches have recently been developed to deal with the fixed effects quantile regression (FE-QR) case, differing primarily (from an applied perspective) in how the fixed effects are viewed.

Consider the following two regression frameworks:

$$\mathbb{E}(Y_{it}|\alpha_i, X_{it}) = \alpha_i + X_{it}\beta \tag{2.3}$$

$$Q_{Y_{it}}(\tau|\alpha_i, X_{it}) = \alpha_i(\tau) + X_{it}\beta(\tau) \tag{2.4}$$

where $Q_{Y_{it}}(\tau|X_{it}, \alpha)$ is the τ^{th} conditional quantile of Y_{it}. Model (2.3) maybe seen to be

for a list of applications to mostly this area of Economics.

obtained from

$$Y_{it} = \alpha_i + X_{it}\beta + \epsilon_{it} \tag{2.5}$$

by imposing the following orthogonality condition: $\mathbb{E}(\epsilon_{it}|X_{it}, \alpha_i) = 0$, while model (2.4) might appear, at a first glance, to be obtained by imposing the following restriction on the model in (2.5): $Q_{\epsilon_{it}}(\tau|X_{it}, \alpha_i) = 0$ (along with the separability assumption on $\alpha_i$and ϵ_{it}). However, the two models are **not** equivalent, except under special conditions, as indicated in detail in Rosen [2009] and we do not further pursue this path here.

In both cases, the α_i[11] are (a) unobservable (b) potentially correlated with the X_{it} and with the residuals of the model (accounting for omitted variables effects). We also maintain the assumption that the X are (conditionally) *exogenous* i.e. $X|\alpha \perp \epsilon$. The $\{\alpha_i\}$ in the equation (2.3) represent "unobserved" heterogeneity while such an interpretation in the case of equation (2.4) is somewhat misleading, since even in the absence of $\{\alpha_i\}$, there is considerable heterogeneity allowed by the model. In fact, the $\{\alpha_i\}$ are not "fixed" in the true sense, since they vary across quantiles, and provide additional flexibility in modeling unobserved heterogeneity (since they vary over the distribution of Y), due to which Harding and Lamarche [2009] term them "Quantile Fixed Effects".

A fundamental issue in both equations is how the $\{\alpha_i\}$ are viewed; as "nuisance" parameters to be estimated or as ancillary parameters aiding identification. In the case of model (2) above, both perspectives yield, from the point of view of identification and estimation of β, simple solutions; identification is straightaway guaranteed under simple orthogonality condition above while estimation is facilitated by noting that the linearity of the expectation operator ensures that the mean of the difference is the difference of the means. In the case of the quantiles, however, conditions on identification of β are not always clear and the

[11] In general, we will not index the parameters α_i and β by τ, except when such an indexing is required, in order to avoid a proliferation of notation, with the understanding that all parameters under consideration are at the same quantile (unless otherwise stated).

"de-meaning" estimation procedure is no longer valid since quantiles of differences are only rarely equal to differences of quantiles[12].

Harding and Lamarche [2009]; Koenker [2004]; Lamarche [2010] pursue the statistics-based approach of treating the $\{\alpha_i\}$ as "nuisance" parameters and focus on consistent estimation while minimizing variance. In more detail, Koenker [2004] views the least squares problem in model (1) in the context of a penalized estimator with an l_2-penalty, the rather special form of the penalty inducing a separation between the estimation of α_i, the fixed effects, and β. Koenker [2004] then proposed both a penalized, with an l_1-penalty[13], and an unpenalized version of FE-QR. The two proposals therein were

$$\min_{\alpha,\beta} \sum_{k-1}^{q} \sum_{j=1}^{n} \sum_{i=1}^{T_i} w_k \rho_{\tau_k}(y_{it} - \alpha_i - X_{it}\beta(\tau_k)) + \lambda \sum_{i=1}^{n} |\alpha_i| \tag{2.6}$$

with the unpenalized estimator obtained by setting $\lambda = 0$, where w_k are the weights on the quantiles. The weighting QR pursued here may be seen as yet another variance reduction device. However holding the $\{\alpha_i\}$ fixed across the quantiles imposes strong restrictions on the type of dependence between α_i and X_i as well as on the omitted variable bias that such effects are typically used to control for in applied studies[14].

This approach follows a rich statistical literature and views the problem in (4) above as one of regularization, since the benefits of an l_1-penalty are well documented in the

[12]To see this in a more precise manner, note that the quantile operator does not commute with the "subtraction" operator; formally, this means that while $\mathbb{E}(Y_{it} - Y_{i,t-1}) = \mathbb{E}(Y_{it}) - \mathbb{E}(Y_{i,t-1})$, $Q_{Y_{it}-Y_{i,t-1}}(\tau) \neq Q_{Y_{i,t}}(\tau) - Q_{Y_{i,t-1}}(\tau)$, where for notational simplicity we do not explicitly indicate that all of these operations are conditional on X_{it}, α_i. This implies, in particular, that "first differencing", which in the case of the conditional mean regression yields a separation in the estimation of the β, is no longer valid, since the quantile of the first difference is not guaranteed to be identical to the difference of the quantiles.

[13]The l-spaces mentioned herein refer to the "little l" norms in functional analysis, the discrete analogue of the L_P-norms. In other words, the l_1-penalty corresponds to the absolute value penalty, the l_2 to the squared penalty.

[14]For instance, viewing the fixed effects as capturing the effect of unobserved, individual specific variables, i.e. if $\alpha_i = Z_i\gamma(\tau)$, restricting them to be invariant across the quantiles implies that the coefficients of such variables are τ-invariant i.e. $\gamma(\tau) = \gamma$.

statistics literature. In effect, the (scalar) l_1-penalty term "shrinks"[15] the estimates of α in order to reduce the variability induced in the estimation of β. While appealing from a theoretical perspective, the major drawbacks in this framework are (a) selection of the shrinkage parameter (b) inference. Both the estimates and the variance matrix are functions of the shrinkage parameters, and while there are results in the case of univariate covariates (Lamarche [2010]), there is no generalization to a multi-variate framework of a method for selection of the shrinkage parameter. An important point to note in the formulation above is that the $\{\alpha_i\}$ are viewed as purely location shift (since they are τ-invariant) .

Kato *et al.* [2010] provide clearer proofs of the asymptotic properties of the unpenalized estimator and investigate the properties of different types of bootstraps for inference in this framework. Further, they also allow the "fixed effects" to vary across quantiles. We therefore follow this approach in our analysis[16].

2.3.3 Identification and Interpretation

There are three important issues to note with the framework above. First, following Powell [2009], it is important to note that the quantiles are defined, in the approaches outlined above, based on each individual unit's residual i.e. based on $\epsilon_{i,t} = Y_{i,t} - \alpha_i - X_{it}\beta$ (the alternative definition being $(\alpha_i + \epsilon_i)$) as a result of which one "loses" the position of an observation in the distribution, compared to the cross-section distribution. However, given that the *effects* identified here (weather on yield) *require* a panel, and do not make sense in a cross-section, it is evident that the definition used here is the most sensible one. This, however, leads to the interpretation of the coefficients being somewhat different from those

[15] There is an extensive statistical literature on the benefits of l_1-penalty and shrinkage. In effect, the main arguments are the statistical and computational benefits offered by this approach. For more details, see Tibshirani [1996], as well as http://www-stat.stanford.edu/∼tibs/lasso.html.

[16] All of the frameworks considered above are subject to the incidental parameter problem (see Lancaster [2000] for an detailed introduction to the topic in econometrics), which is "solved" in all cases above by allowing *both* n and T to diverge at specific rates.

in the purely cross-sectional case[17].

Second, if year fixed effects are not included in the above framework, then, as pointed out in Powell [2009], the resulting τ^{th} quantile corresponds to different parts of the distribution for different years, a possibly undesirable artifact in applications. Third, unlike in the case of the linear panel data models (where identification and consistency follows from a simple orthogonality condition between X and ϵ), conditions for identification and consistency of β involve high level conditions on the marginal density of the residuals evaluated at zero and on moment conditions on X_i (Assumption A3 in Kato *et al.* [2010]), due to which they have no interpretable or intuitive content.

2.3.4 Covariance Matrix Estimation

For the unpenalized case, Koenker [2004] derives the asymptotic variance of the estimator and points out that bootstrap inference maybe more accurate (another cause for concern is that the asymptotic variance requires T to grow "fast enough", compared to N; in particular, $\frac{N^a}{T} \to 0$ for some $a > 0$). Kato *et al.* [2010], for the same unpenalized model, also point out that asymptotics require, in their case, T to grown faster than N^2, a clearly restrictive condition, and they recommend the cross-sectional bootstrap as being the better approach, after comparing both cross-sectional and time bootstraps for their coverage accuracy. We note however that their comparisons are for coverage for *individual coefficients* only.

There is a large literature on the use of the bootstrap for confidence interval estimation in a cross-section QR setting (see for instance Buchinsky [1995]; Fitzenberger *et al.* [2002]; Fitzenberger [1998] among others). For the panel data case, there are three possible basic version of the bootstrap, which, following Kapetanios [2008], are termed "Time" bootstrap,

[17] It is, in this context, important to note that if the use of a panel framework is for the purposes of identifying the effects of a variable which is believed to be endogenous in the cross-section case (such as for instance in the hedonic case), then the framework outlined above is likely inappropriate, and a better suited one is the GMM approach outlined in Powell [2009]

with resampling across the time dimension, "Cross-section", with resampling across the "individual unit" dimension and a combination. We note that all three versions correspond to the "paired" or the "x-y" bootstrap, involving bootstrapping the data matrices rather than the residuals[18]. We use the simple time bootstrap for inference, which accounts for the unknown spatial dependence but not for time dependence. Details of estimation of the standard errors (and therefore, the t-statistic) are provided in the Appendix B.

2.3.5 Extension to semi-parametric quantile regression

We outline an extension of the above linear conditional QR-FE estimation strategy to semi-parametric settings. Consider the linear FE-QR estimation detailed above:

$$Q_{Y_{it}}(\tau|\alpha_i, X_{it}) = \alpha_i(\tau) + X_{it}\beta(\tau) \tag{2.7}$$

and observe that, once a particular estimation framework is settled upon, the problem above is in essence a straightforward linear QR problem (with more involved conditions for consistency). Consider instead the more general semi-parametric conditional quantile estimation problem:

$$Q_{Y_{it}}(\tau|\alpha_i, X_{it}) = \alpha_i(\tau) + X^1_{i,t}\gamma + g(Z_{it}) \tag{2.8}$$

where Z_{it} is (for now) a univariate variable to be modeled non-parametrically. There exist two possible approaches to non-parametric modeling of quantiles (see Koenker [2005, Chapter 7] for a survey), the local-polynomial based one, and a penalized spline version.

[18] In more detail, denoting $t_{i_1}, \ldots, t_{i_T}$ the resampled time indices (assuming for now a balanced panel data), time resampling involves forming the full bootstrapped data, $Z^* = (Y^*, X^*)$ where $Y^* = (Y_{t_{i_1}}, Y_{t_{i_2}}, \ldots, Y_{t_{i_T}})$ and $X^* = (X_{t_{i_1}}, \ldots, X_{t_{i_T}})$ where Y_t, X_t denote respectively the $N \times 1$ and $N \times K$ matrix of observations for all N individual units. Cross-sectional resampling is entirely analogous, with $I_{j_1}, I_{j_2}, \ldots, I_{j_N}$ denoting the resampled individual unit indices, $Y^* = (Y_{I_{j_1}}, Y_{I_{j_2}}, \ldots, Y_{I_{j)N}})$, $X^* = (X_{I_{j_1}}, \ldots, X_{I_{j_N}})$ with $Y_{I_{j_n}}$ being the resampled $T \times 1$ vector of observations on individual I_{j_n} for all T time periods and $X_{I_{j_n}}$ the $T \times K$ matrix of covariates for individual unit I_{j_n}

We note however, that even without these options, the function $g(Z_{it})$ may always be approximated by an orthogonal polynomial of a suitable order, and the problem is one of estimation of the order of the polynomial. Inference is even more problematic in this setting; however, since consistency is likely guaranteed (under possibly additional conditions), a bootstrap approach maybe easily used in this case (and this is the approach pursued here).

The alternative approach, of penalized splines, is a different variant of the same approach, in that the scalar penalty term must be chosen, generally through means of AIC-like measures (see Koenker [2010] for details and an application); we briefly pursue this approach here, more as a robustness check, but note that in most non-parametric QR cases, inference is generally problematic and computationally expensive, and this is especially true in the FE case.

2.3.6 Model Specification

The basic model we consider is of the following form, for a given crop:

$$\begin{aligned} y_{it} &= \alpha_i + \beta_1 GDD_{it} + \beta_2 GDD_{it}^2 + \beta_3 Precip_{it} + \beta_4 Precip_{it}^2 \\ &\quad + \sum_{k=1}^{K} \sum_{j=1}^{J} \beta_{5+k,,j} Trend^k * R_j \end{aligned} \tag{2.9}$$

where $y_{it}$is the log of yield, $R_j$is a region indicator (we use 5 regions, North, West, South, East and Central with "North" being the omitted category), $K = 3$ is the order of the time trend. A second possible specification where we use, in addition, monsoon onset at the district level and the number of dry spells for the summer growing season, is

$$\begin{aligned} y_{it} &= \alpha_i + \beta_1 GDD_{it} + \beta_2 GDD_{it}^2 + \beta_3 Precip_{it} + \beta_4 Precip_{it}^2 \\ &\quad + \sum_{k}^{K} \sum_{j}^{J} \beta_{5+j,k} Trend^k * R_j + \beta_{5+J+1} Onset_{it} + \beta_{5+J+2} dryspells_{it} \end{aligned} \tag{2.10}$$

The third specification is where we use monthly precipitation data instead of using seasonal totals as above.

$$\begin{aligned} y_{it} &= \sum_{\kappa=1}^{K_1} precip_{kit}\beta_k + GDD_{it}\beta_{K+1} + GDD_{it}^2\beta_{K+2} + \sum_{j=1}^{K} precip_{kit}^2\delta_j \\ &\quad + \sum_{j}^{J} \sum_{k}^{K} \gamma_{k,j} Trend * R_j + \nu Onset_{it} + \zeta dryspells_{it} \end{aligned} \tag{2.11}$$

with $K_1 = 6$ being the number of months in the winter (rabi) growing season, and $K_1 = 4$ the number of months in the summer growing season (kharif), with onset and dry spells included for the summer growing season only. Another covariate added to many regressions is (where available) the ratio of irrigated area to un-irrigated area at the district level. Finally, for the winter season which spans 6 months, instead of including rainfall for all 6 months separately, in one specification, we aggregate them into 3 separate periods, early, middle and late season rainfall[19].

[19] For instance, there is generally little rainfall during the months of Feb-March and therefore including them separately provides little additional knowledge.

2.4 Data and summary statistics

2.4.1 Agricultural Data

We use district-level data on agricultural outcomes from the commercially available *Indian Harvest* database from the **CMIE**. The database includes crop-specific data on district-level yield, area sown, farm harvest price and sparse input data, in addition to irrigated area by crop. Data on inputs are very unreliable, and we do not make use of them. The dataset nominally spans the time period 1950-2006 but data for a large proportion of the districts are available only from 1970 onwards and we therefore use data from 1971. The dataset covers most of the major crops (including what are termed "cash crops" but not plantation crops such as "tea", "coffee" and "rubber") and all the states and districts. The present analysis is restricted to the major cereals *rice* and *wheat*. The data are essentially a compilation of the official statistics reported by the state governments and the Ministry of Agriculture of the Govt. of India.

The CMIE's Indian Harvest database has only been available since 2003. Previous research for India on this topic used a dataset compiled by the World Bank (WB) (Sanghi *et al.* [1998]) which pre-dates the CMIE's database and is essentially similar. The main differences stem in geographical coverage and homogeneity of district definition. The WB dataset covers a large proportion of the agriculturally important states while the CMIE database covers all states. Further, data on the large and agriculturally important states of Bihar and West Bengal are mostly missing in the extension by Duflo and Pande [2007].

Indian district boundaries change periodically, with larger districts generally split into smaller ones (sometime combinations of districts are split into larger numbers but this is rare), creating difficulties for defining a district as the unit of observation. Further, no year-by-year spatial definition of the districts are available, and it is not always clear what year a district was changed; we therefore take the year in which data for a district appears

in the agricultural data as the "year" in which the said district was created. Yet another difficulty lies in that the "mother" district, from which the new district is created, exists and has the same name as before. However, in many cases, the agricultural data base indicates that a change has occurred[20], and this allows a largely[21] comprehensive accounting of (a) which new districts are created (b) which existing districts have altered and at what point of time.

There are two ways to proceed to obtain uniform district data sets. The first approach, pursued in Duflo and Pande [2007] (an extension of the earlier dataset compiled by Sanghi *et al.* [1998]), is one in which districts were defined based on the most primitive definition and are therefore consistent in time. The second approach, and the one we pursue, is to (a) track when new districts are created (b) treat the new districts as being distinct from the original ones[22]. We argue that this is a better indicator of the districts for the following reason: in general, districts are altered for ease of administration and/or for historical reasons. We anticipate the two districts to have different characteristics (in addition to the obvious one of different area), and therefore, separate fixed effects for each is more appropriate when the fixed effects are viewed as accounting for "unobserved heterogeneity". In order to obtain geographical co-ordinates for each district (primarily the district center), we make use of GIS maps of India for two time periods, 1991 and 2001. This allow us to keep track of how district centroids change with time, following the approach outlined above in treating districts which undergo any changes as separate districts (although most districts are already small enough that even after a change, their centroids are either almost

[20]The timing of the change is generally restricted to the census years, due to the nature of reporting of data.

[21]This is "largely" but not completely comprehensive since it is not clear if all districts which have altered even slightly have indicators in the database but this issue is likely to be minor enough to be ignored.

[22]Thus, if district "X" was split in year "Y" into districts "X(1)" and "Z" (the "(1)" is simply an indicator–generally a phrase such as "up to 1981" etc–allowing one to distinguish between the pre- and -post districts), then we treat districts "X", "X1" and "Z" as *all distinct*, with district "X" being in existence prior to year "Y", and districts "X1" and "Z" existing post year "Y". This approach of course leads to the number of districts being different at different time points (generally, the census years).

unchanged or change at most by half a degree)[23].

2.4.2 Weather Data

2.4.2.1 Temperature Data

Schlenker and Roberts [2006, 2009] have illustrated the superiority of using fine-scale temperature measures to account for the cumulative effects of variations in temperature on crop yield. Much of this research has been confined to the developed countries (an exception is Guiteras [2008]) due to lack of availability of such data for the developing nation setting. Given the paucity of station-level temperature and rainfall data for India, prior work (Guiteras [2008]) makes use of a corrected reanalysis gridded (at the $1^\circ \times 1^\circ$ resolution) dataset, with 6-hourly measurements of temperature from 1949-2000. This study makes use of gridded ($1^\circ \times 1^\circ$ resolution) daily rainfall and temperature products recently created and made available by the IMD (Srivastava *et al.* [2009]). To our knowledge, this is the first paper in Economics to make use of this dataset, spanning 1969-2005.

To create daily district-level data from gridded, we use a weighted (using standard weights in the Climate literature, the inverse square root of distance) average of data from all grids which are within a 100 KM radius[24] from the district centroid. Due to the nature of agricultural policy, yields across space are expected to be related (especially within a state and region); further, the mode of creation of the weather variables (gridding) induces additional dependence across space, due to which dependence of the regression residuals are likely to be dependent across space. Following recent literature on bootstrapping the variance estimators for panel data with spatial dependence (Gonçalves [2008]), we illustrate

[23]The approach in the CMIE database to new states is to treat these states and their districts to have existed for the entire period and to simply relabel the districts originally in another state to have been in the new state. For instance, districts which were in Madhya Pradesh and were later a part of Chhatisgarh are treated as being, for all time periods, in Chhatisgarh. We follow this convention since it is also more sensible when reporting effects at the state level.

[24]Alternative radii of up to 200 km yielded very similar seasonal summary statistics.

a simple bootstrap approach which accounts for this, even in the relatively complicated estimation framework employed here.

2.4.2.2 Precipitation

For precipitation, we again make use of a gridded daily rainfall data set product, spanning 1951-2003, developed by the IMD (Rajeevan *et al.* [2005]). To our knowledge, this is the first paper in Economics to make use of this dataset. Creation of district-level data from gridded proceeds in the same way as for temperature. The significance of using actual data, in comparison to reanalysis data, for India is illustrated elsewhere (and is available upon request) but the main differences are that the fidelity to actual rainfall varies with reanalysis datasets, and the NCC dataset in particular appears poorly constrained due to the lack of actual data for India.

2.4.3 Summary statistics

Summary statistics of the key variables used, by region, are presented in Table 2.1 for wheat and Table 2.2 for rice. Three points are noteworthy about Table 2.1: the predominance of the Northern region in wheat production, its high yields and the very high irrigation ratios. Turning now to rice, it is evident that the Eastern region dominates in rice production, followed by the Southern and Northern regions; in addition, there is a significant difference in rainfall between the Eastern regions and the rest of India. Finally, the relatively late onset of the monsoon is evident in the Northern and Western regions[25].

[25]The CMIE data base, and in general agricultural statistics in India, reports irrigated area by crop and year, instead of by season. However there are, for certain districts in South Eastern India, two rice growing seasons, due to which we have been unable to establish the extent of irrigation during the kharif season. We therefore do not have potentially important information regarding irrigation in the kharif season.

Table 2.1: Descriptive statistics of regression sample for wheat (*rabi* season). Mean coefficients reported, with standard deviations in parenthesis.

	Central	North	South	West	Total
yield(kg/hectare)	1054.8	2056.1	786.1	1313.0	1551.2
	(506.5)	(829.5)	(438.5)	(592.0)	(849.6)
production(000's of ton)	80.24	335.4	7.089	84.51	193.5
	(63.28)	(297.1)	(15.09)	(106.3)	(251.9)
Seasonal rainfall (mm)	88.13	105.0	185.0	67.69	102.7
	(68.31)	(73.48)	(101.9)	(65.68)	(81.44)
growing degree days (deg Celsius)	2808.2	2444.1	3383.1	2845.5	2708.1
	(217.9)	(311.5)	(164.4)	(351.3)	(417.3)
harmful degree days(deg Celsius)	8.499	4.146	32.16	15.04	10.43
	(8.898)	(5.702)	(23.50)	(14.92)	(14.70)
Average seasonal temp (deg Celsius)	21.44	19.41	24.74	21.68	20.90
	(1.233)	(1.727)	(0.993)	(1.985)	(2.353)
irr_ratio	0.407	0.849	0.596	0.673	0.692
	(0.293)	(0.180)	(0.334)	(0.274)	(0.301)
Observations	911	2053	486	893	4343
Number of district	45	97	28	52	222

Table 2.2: Descriptive statistics of regression sample for rice (*kharif* season). Mean coefficients reported, with standard deviations in parenthesis.

	East	North	South	West	Total
yield (kg/hectare)	1206.4	1593.6	1961.8	1098.0	1523.4
	(462.8)	(863.6)	(645.0)	(562.7)	(757.6)
production (000's of ton)	310.0	161.0	197.6	39.46	170.8
	(248.5)	(174.4)	(217.5)	(75.34)	(207.2)
Seasonal rainfall (mm)	1346.2	782.4	844.7	801.4	904.6
	(546.9)	(297.7)	(707.0)	(451.2)	(562.6)
growing degree days (deg Celsius)	2460.8	2625.7	2320.5	2477.4	2472.1
	(147.7)	(120.9)	(293.9)	(167.5)	(233.0)
gdd lower threshold	2456.3	2598.7	2315.2	2462.3	2458.3
	(145.3)	(107.4)	(289.2)	(155.3)	(222.7)
harmful gdd (deg Celsius)	0.754	11.01	1.383	5.968	5.194
	(2.716)	(13.37)	(4.285)	(11.14)	(10.32)
hgdd lower threshold	5.161	38.03	6.654	21.06	19.05
	(9.274)	(29.74)	(12.74)	(28.11)	(26.37)
Day of monsoon onset (days from April 1)	29.41	72.20	47.97	76.63	58.43
	(23.57)	(21.41)	(31.92)	(19.13)	(30.56)
dry spells (number of days)	3.964	5.067	6	5.014	5.139
	(1.786)	(1.575)	(1.858)	(1.536)	(1.826)
Average seasonal temp (deg Celsius)	28.18	29.62	27.04	28.37	28.31
	(1.218)	(1.056)	(2.420)	(1.426)	(1.950)
Number of districts	36	66	60	56	218
Observations	880	1497	1490	1127	4994

2.5 Regression results

As indicated earlier, for each of rice and wheat, two sets of regressions are carried out for each specification of GDD: using seasonal rainfall and using monthly rainfall. For the sake of brevity, and since in our climate change specifications we do not use monthly rainfall changes, we report and discuss results using only seasonal rainfall[26]. The regression strategy is as follows: interest centers on the estimation of the τ^{th} quantile of *log(yield)*, as in equation (2.9) for the parametric case and equation (2.8) for the non-parametric case. The actual equations estimated, however, are (2.9), for the rabi season, and (2.10) and (2.11) for kharif, substituted into equation (2.6) (with equal weights and no penalty, as indicated earlier).

We follow much of the literature (see Fitzenberger *et al.* [2002]) and chose the first three quartiles, along with the $90^{th}\%tile$: $\tau = (0.25, 0.5, 0.75, 0.9)$ for estimation. All regressions include region-specific cubic time trends whose coefficients, along with those of individual fixed effects, are not reported since they are not of intrinsic interest. We avoid the use of year fixed effects for two reasons, to better capture region-specific heterogeneity and to avoid numerical issues with the linear programming problem[27]. An undesirable consequence of this choice is that the interpretation of the coefficients at each quantile are somewhat different from those in a pure cross-section, as indicated earlier. As indicated above, both asymptotic and (time) bootstrap t-statistics are reported.

2.5.1 Wheat

Results of estimation are provided in Tables (2.3) and (2.4). Given the log-linear nature of the regressions, we may interpret the coefficients as approximately corresponding percentage

[26] Nearly identical results were obtained when monthly (in case of kharif) and bi-monthly (in case of rabi) rainfall were used instead.

[27] It proved difficult to estimate the year fixed effects, given the unbalanced nature of the regression sample and the fact that both individual fixed effects and year fixed effects must be estimated. For the same reason, it is even more infeasible in this approach to estimate region-specific or district-specific year fixed effects.

Table 2.3: Quantile and linear panel regression with fixed-effects for wheat, using seasonal rainfall. All regressions include region-specific cubic time trends and district fixed-effects whose coefficients are not reported. T-statistics for the test $H_0 : \beta = 0$ based on the asymptotic variance matrix ({}) and time bootstrap for panel data ([]) are reported.

	tau=0.25	tau=0.5	tau=0.75	tau=0.9	Linear Model
GDD	2.05E-03	1.94E-03	1.51E-03	1.41E-03	2.04E-03
	{4.106}	{5.08}	{4.86}	{4.034}	{6.626}
	[3.862]	[4.802]	[4.633]	[4.621]	
GDD sqr	-4.79E-07	-4.21E-07	-3.26E-07	-3.32E-07	-4.53E-07
	{-5.133}	{-5.941}	{-5.367}	{-4.493}	{-7.98}
	[-4.841]	[-5.632]	[-5.221]	[-5.303]	
Seasonal Rain	5.03E-04	2.91E-04	6.41E-04	3.69E-04	4.95E-04
	{2.19}	{1.88}	{3.685}	{1.739}	{3.3}
	[2.446]	[1.69]	[3.233]	[1.976]	
Seasonal rain sqr	-1.26E-06	-9.97E-07	-1.74E-06	-4.09E-07	-1.28E-06
	{-1.772}	{-2.483}	{-3.168}	{-0.5686}	{-3.15}
	[-2.032]	[-2.143]	[-2.698]	[-0.6852]	
hgdd	-7.58E-04	-4.10E-04	5.45E-05	1.27E-03	3.79E-04
	{-0.827}	{-0.5346}	{0.07862}	{1.419}	{0.654}
	[-0.7863]	[-0.5426]	[0.07665]	[1.627]	
Irr_ratio	1.00E+00	9.75E-01	9.27E-01	7.75E-01	7.95E-01
	{14.83}	{15.34}	{12.88}	{8.697}	{22.92}
	[15.25]	[15.72]	[14.59]	[9.505]	

Table 2.4: Same as table(2.3) but using orthogonal polynomials (of degree 4) for gdd.

	tau=0.25	**tau=0.5**	**tau=0.75**	**tau=0.9**	**Linear Model**
Seasonal Rain	5.08E-04	2.38E-04	5.67E-04	3.96E-04	4.84E-04
	{2.232}	{1.544}	{3.183}	{1.874}	{3.23}
	[2.519]	[1.42]	[2.9]	[2.124]	
Seasonal rain sqr	-1.29E-06	-8.63E-07	-1.52E-06	-4.82E-07	-1.26E-06
	{-1.851}	{-2.19}	{-2.686}	{-0.6792}	{-3.09}
	[-2.118]	[-1.902]	[-2.447]	[-0.7924]	
hgdd	-8.45E-04	-4.62E-04	-6.85E-05	1.28E-03	3.13E-04
	{-0.9137}	{-0.6013}	{-0.09465}	{1.436}	{0.53}
	[-0.8716]	[-0.597]	[-0.0917]	[1.647]	
Irr_ratio	1.01E+00	9.71E-01	9.22E-01	7.76E-01	7.95E-01
	{14.99}	{15.33}	{12.32}	{8.78}	{22.92}
	[15.32]	[15.75]	[14.26]	[9.397]	

changes. Turning first to effects of temperature, two conclusions are evident. First, there appears to be a inverted U-shape relationship between yield and temperature, indicating that temperature up to a certain point is beneficial, and harmful thereafter. This is in keeping with the results obtained in Schlenker and Roberts [2009] and in Guiteras [2008], using very different methods. A second, and more surprising, one is that the differences across quantiles of yield of the inflection point is minimal, indicating a significant degree of uniformity in impacts across differing crop and ecology types (controlling for irrigation and rainfall). Rainfall also appears to have a similar relationship to yield as does temperature and interestingly, the coefficients across different quantiles is very different (for instance, in absolute terms, the coefficients in many instances differ by up to 50%), reflecting the high spatio-temporal variability in winter season rainfall (evident in Table 2.1).

Yet another interesting point is that while Harmful Degree Days appear to have (for the most part) the correct sign, they are never significant, indicating that, when accounting for a richer interaction between temperature and rainfall (as the log specification does[28]),

[28]We note that the use of a log specification implies some amount of substitutability between heat and

the additively separable nature of Harmful Degree Days yields little additional information. Irrigation, unsurprisingly, is highly significant and positive at all quantiles, with a possibly smaller coefficient at the 90^{th}percentile. Results using bi-monthly rainfall totals are almost identical (and are available upon request), although given the highly variable and spatially concentrated nature of rainfall, it is difficult to interpret the coefficients on rainfall.

Finally, in order to assess if the results are influenced by the quadratic specification for GDD, we use two, very different, non-parametric approaches; first, we use an orthogonal polynomial expansion for growing degree days, using polynomials of degree 4, and second, we use the penalized spline approach outlined above. The coefficients for other variables, using either approach, appeared to be identical (we do not therefore report coefficients estimated with the spline specification), as illustrated for the case of orthogonal polynomials in Table 2.4. The estimated effect of GDD on yields using non-parametric splines, in Figure 2.1, indicates very clearly an initial increase, followed by a subsequent decrease (the inflection point also appears very close to the ones obtained using simple quadratic specification above). Thus, the results are *robust* to functional form assumptions[29].

2.5.2 Rice

In the case of rice, GDD appear to have a U-shaped relationship, quite contrary to intuition and agronomic belief, as indicated in Table 2.5. Further, this behavior is consistent across quantiles, indicating that the relationship is global (in the distribution function). More surprisingly, a very similar relationship appears to hold for rainfall, and rainfall is seen to be significant at all but the lowest quantile; similar is the case with onset day and dry spells. In order to assess if the results are impacted by the simple, quadratic functional form assumptions, both orthogonal polynomials and penalized splines were used to model

rainfall, which is realistic. It also limits the interactions to be multiplicative.

[29]More surprisingly, the *direction* of impact of temperature and rainfall are consistent even if very crude measures of temperature, such as average daily district temperature, are used. See for instance Table 2.10

Table 2.5: Quantile and linear panel regression with fixed-effects for rice, using seasonal rainfall data. All regressions include region-specific cubic time trends and district fixed-effects whose coefficients are not reported. T-statistics for the test $H_0 : \beta = 0$ based on the asymptotic variance matrix ({}) and time bootstrap for panel data ([]) are reported.

	tau=0.25	**tau=0.5**	**tau=0.75**	**tau=0.9**	**Linear Model**
GDD	-4.35E-03	-3.51E-03	-2.59E-03	-2.33E-03	-3.81E-03
	{-5.029}	{-4.101}	{-3.171}	{-2.46}	{-3.09}
	[-4.744]	[-4.269]	[-3.357]	[-2.601]	
GDD sqr	9.58E-07	7.41E-07	5.17E-07	4.53E-07	8.35E-07
	{5.6}	{4.346}	{3.095}	{2.339}	{3.45}
	[5.186]	[4.515]	[3.335]	[2.53]	
Seasonal Rain	-8.26E-05	-3.36E-05	3.73E-06	-5.32E-07	-2.10E-04
	{-1.763}	{-0.8878}	{0.103}	{-0.01362}	{-4.47}
	[-1.625]	[-0.8545]	[0.1049]	[-0.01319]	
Seasonal rain sqr	1.57E-08	3.99E-09	-4.14E-09	1.05E-09	4.47E-08
	{1.461}	{0.4743}	{-0.5071}	{0.09544}	{3.09}
	[1.349]	[0.4675]	[-0.4921]	[0.09404]	
onsetday	-2.13E-04	-2.37E-04	-5.02E-04	-4.46E-04	-1.22E-04
	{-0.8656}	{-1.175}	{-2.605}	{-2.267}	{-0.43}
	[-0.883]	[-1.184]	[-2.676]	[-2.335]	
dryspells	-5.30E-03	-1.77E-03	-3.62E-03	-3.96E-03	-1.14E-02
	{-1.258}	{-0.563}	{-1.208}	{-1.292}	{-2.78}
	[-1.182]	[-0.5704]	[-1.258]	[-1.354]	

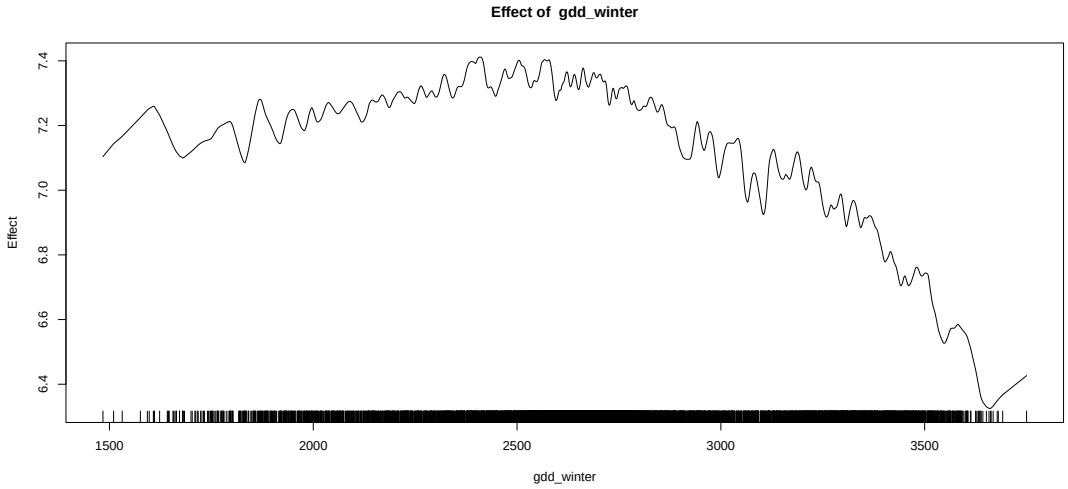


Figure 2.1: Non-parametric (penalized spline) estimation of the effect of Growing Degree Days on wheat yields (at the median).

Table 2.6:]
FE-QR for rice using seasonal rainfall data and orthogonal polynomials of degree 4 for gdd (coefficients not shown). Other conventions same as for Table (2.5)

	tau=0.25	**tau=0.5**	**tau=0.75**	**tau=0.9**	**Linear Model**
Seasonal Rain	-9.01E-05	-4.22E-05	-2.34E-06	-1.06E-05	-2.16E-04
	{-1.939}	{-1.11}	{-0.06455}	{-0.2742}	{-4.59}
	[-0.3732]	[-0.2109]	[-0.01245]	[-0.05542]	
Seasonal rain sqr	1.77E-08	5.76E-09	-2.20E-09	1.56E-09	4.57E-08
	{1.695}	{0.6882}	{-0.2674}	{0.145}	{3.162}
	[3.95e-06]	[1.855e-06]	[-7.634e-07]	[5.34e-07]	
onsetday	-1.53E-04	-2.73E-04	-4.80E-04	-4.19E-04	-1.17E-04
	{-0.63}	{-1.363}	{-2.49}	{-2.106}	{-0.42}
	[-0.01368]	[-0.0368]	[-0.05795]	[-0.04873]	
dryspells	-6.09E-03	-1.69E-03	-3.89E-03	-4.09E-03	-1.16E-02
	{-1.45}	{-0.54}	{-1.3}	{-1.326}	{-2.85}
	[-7.238]	[-3.03]	[-6.407]	[-6.313]	

GDD.

Results from Table 2.6 indicate the quadratic specification for GDD performs poorly and tends to inflate variance of the remaining parameters, while both alternative specifications perform much better. However, for brevity, we report only results using the orthogonal polynomial specification, in Table 2.6. The impact of rainfall on yield appears very similar across the quadratic and orthogonal polynomial specification, in being significant at only the first quartile; the impact of onset day too appears rarely significant in both specifications. Dry spells however appear much more precisely estimated, and are, as anticipated, significantly more harmful at the lower quantiles than at the higher (the coefficients differ by about *30%*).

Further, the fully non-parametric approach to estimating the effect of GDD, as indicated in Figure 2.2, reinforces the robustness of an inverted U-shape relationship. Thus, it is clear that the relationship outlined, and the lack of significance, are robust to various functional forms. These results are unaltered when monthly rainfall is used instead of seasonal rainfall. Yet another concern could be the omission of irrigation, which is of some importance for many regions, especially the Northern region. Given the lack of reliable data on irrigation at the district level for rice by season, it is not possible to directly address this concern. However, we argue that it is very plausible that the possible bias due to omission of irrigation is unlikely to drive our results, by noting that rainfall and onset days are already not significant; thus, the concern that these coefficients are larger (or significant) due to the omission of irrigation is no longer as important.

Intriguingly, however, recent results using farm level data on irrigated agriculture (Welch *et al.* [2010]) indicate that maximum temperatures are *beneficial* for rice crops, despite agronomic evidence to the contrary. While the study had only one site in India, in Tamil Nadu, with maximum temperature significantly lower than those in say, North-Western India, the impact of the interaction of high temperatures with irrigation may well explain

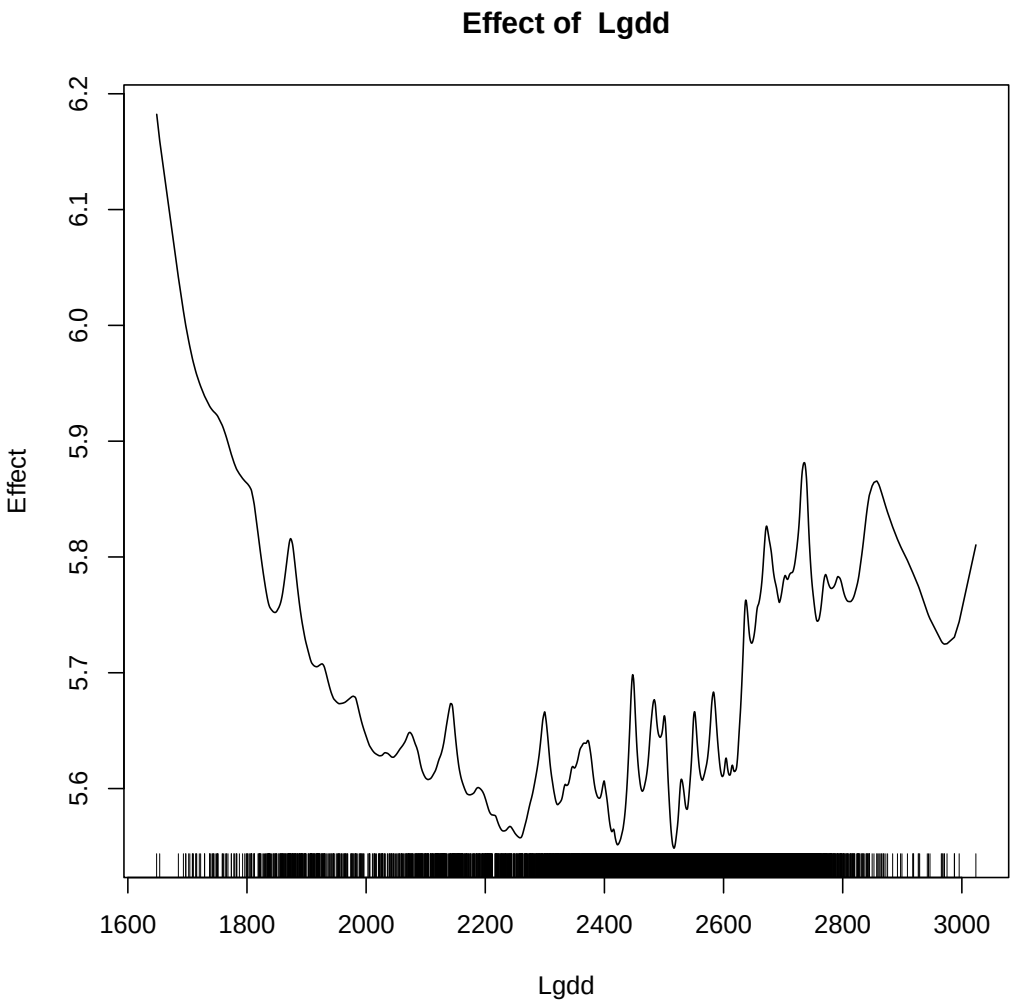


Figure 2.2: Non-parametric (penalized spline) estimation of the effect of growing degree days on rice yields (at the median).

a part of the results obtained here[30].

2.5.3 Comparison with the standard, conditional mean approach

In order to illustrate the benefits of the current approach, consider the results of using an identical specification in a standard, linear panel data framework, results for which are in the final columns of Table 2.3 to 2.6. Since the comparison is identical when using orthogonal polynomials, we focus on the simple, quadratic-in-gdd specification.

Consider first, the case of wheat, in Table 2.3, where magnitudes of differences in coefficients across quantiles are smaller than those for rice. Two points are evident from a perusal of the results: (a) results of the mean regression are very similar to those of the quantiles when there is little difference in coefficients across the quantiles, as for the case of *GDD* (b) when there are substantial differences across quantiles (as in the case of irrigation and rainfall), mean regression provides a very incomplete characterization of the conditional distribution. For instance, in the case of irrigation ratio, the coefficient in the mean regression is 7.95, which is very close to the impact of regression at the 90^{th} percentile, and is *substantially* (about *25%*) *smaller* than at the first quartile.

These differences are even larger in the case of rice (Table 2.5). Consider for instance the coefficient on either onset day or dry spells; these coefficients are, in many instances, substantially (an order of magnitude, in the case of dryspells) different. It is therefore clear that substantial heterogeneity in relationships exist at different quantiles and that the mean regression framework, in this case, provides a very incomplete characterization of the

[30] We also indicate that the results obtained here do not require the use of the linear GDD specification. For instance, the results in Table 2.11 indicate that even with a very crude measure of temperature, district average seasonal temperature, the *direction* of impact of temperature and rainfall are consistent with that obtained using the GDD measure. Very similar results are obtained using GDD computed with a presumed daily distribution of temperature, approximated using a sine-curve by the Baskerville-Emin method, which uses the minimum and maximum temperature, instead of just using the average temperature, as the linear GDD method uses. Therefore, we do not report separately regression results using this method of GDD computation (results are available upon request).

conditional distribution of yield. The substantial differences in impacts using these two frameworks are fully illustrated below.

2.5.4 Climate Change scenarios

Climate change scenarios for econometric studies of impacts on agriculture has focused on GCM projections, for the most part, the Hadley Center models (Deschenes and Greenstone [2007]; Guiteras [2008]; Schlenker and Roberts [2009]). However, for the US, Schlenker and Roberts [2009] used downscaled GCM projections while Guiteras [2008] uses the coarser GCM daily output. GCM daily projections, however, are known to have significant biases which make their use for agricultural purposes quite tricky (Ines *et al.* [2010]). Further, the coarseness of the GCM output tends to lead to a large degree of spatial smoothing, a very undesirable feature. While we hope to report on downscaling GCM output for India and their applications in a future iteration, we will use simpler scenarios for the future at present.

Consistent with IPCC projections and with prior literature on crop modeling for India, we consider the following scenarios for kharif (summer growing) season: uniform increases in temperature over the growing season of 1° and $2^{\circ}C$, coupled with uniform increases in rainfall of 10%. It is clear that, using this approach, the increases in rainfall are significantly overstated (since model projections expect *reduction* in rainfall over certain regions of India, especially the Northwest). Further, many GCM's also project increase in monsoon rainfall variability, which we interpret as increases in probability of late monsoon (by a uniform 10%) and increases in dry spells (again by a uniform 10%), given model projections of more intense rainfall coupled with reduction in rainfall frequency. However, for the results below, we do not apply changes in sub-seasonal characteristics, in order to minimize clutter (and since modest changes in these variables do not substantially affect yield).

For the *rabi* (winter growing) season, there have been very few studies on possible

impacts; however, most indications are for increased minimum (therefore average) temperatures, and little changes in rainfall. We translate this, in line with prior crop modeling efforts, into uniform 0.5° and $1^{\circ}C$ increase in temperature, and an unchanged rainfall distribution. We emphasize that these scenarios are consistent with those used in the crop modeling literature (Mall *et al.* [2006a,b]) as well as prior econometric approaches (Guiteras [2008]; Sanghi *et al.* [1998]).

2.5.5 Climate Change impacts

The impacts of climate change are computed, at each quantile, as the differences in predicted values under current and future weather. In more detail, denoting with superscript f *future* realizations of weather, we compute

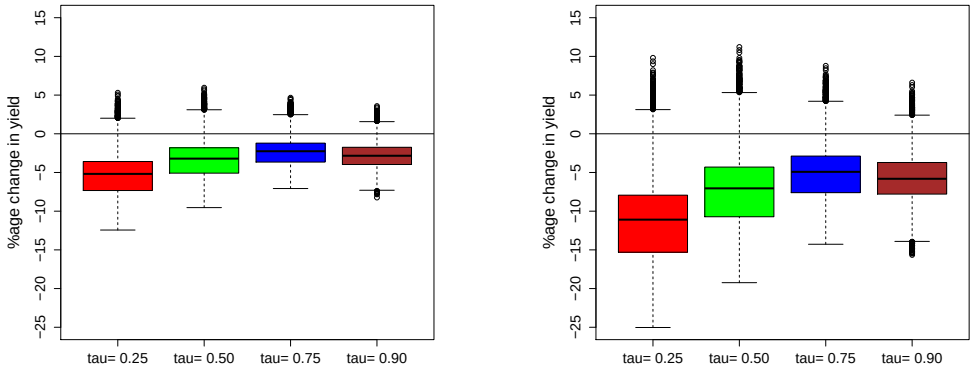
$$\Delta Y_{it}(\tau) = \widehat{Y_{it}^{f}(\tau)} - \widehat{Y_{it}(\tau)} \tag{2.12}$$

as the predicted change. In other words, predicted values are obtained, using coefficients estimated under current weather, for both current and future weather and the difference in predicted values is interpreted as the impact of change in climate. We present the full distribution of predicted changes, instead of focusing on summary changes (such as changes at the mean of the covariates) in order to assess the full range of changes predicted. The distribution of the changes are summarized using box plots.

2.5.5.1 Wheat

There are two possible scenarios for winter, as indicated above. Given the signs of the coefficients, one anticipates significant losses in yield under increased warming. The results in Table 2.7 indicate that increasing temperatures indeed are harmful. Further, the magnitude of losses are decreasing in quantiles, for every scenario considered, indicating a significant degree of heterogeneity in the direction (not magnitude) of impact across a

range of growing conditions. Finally, median losses range from 2 to 5% in the relatively benign 0.5^0C temperature increase scenario to a substantial $5-10\%$ in the 1^0C warming scenario, with many districts losing as much as 22% of yield. Figure 2.3 provides a visual description of the magnitude of losses; for scenario 2, losses rise to as high as 20% for the lower percentiles, indicating significant losses for the already *lower productive* districts[31].



(a) Scenario 1

(b) Scenario 2

Figure 2.3: Box plots of changes in wheat yield under scenarios indicated, for various quantiles.Scenarios are same as in Table (4)

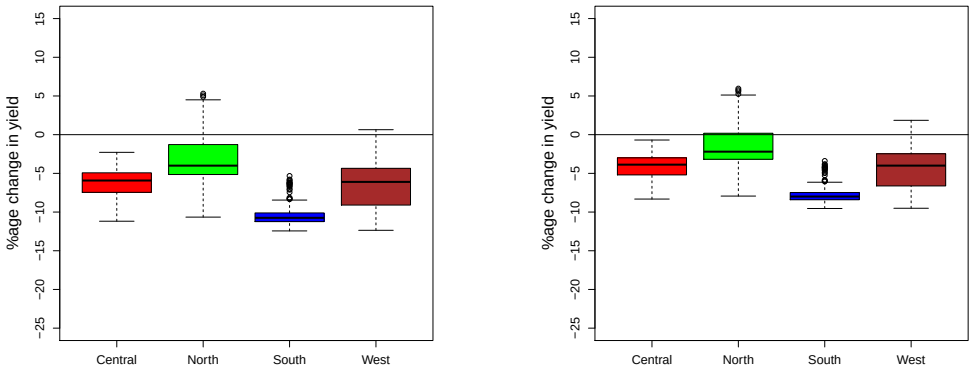
Turning now to assessing regional impacts, we note that the Southern region, already quite hot, has the highest losses (although of less importance in wheat production nationally) in both scenarios, as indicated in Figure 2.4 (we do not include the figure for regional losses under scenario 2, since the regional distribution is identical). However, at all percentiles, Western and Central regions, regions of already high poverty levels and substantial wheat

[31]The interpretation of the coefficients are somewhat different from the cross-section case, since the quantiles are defined for i and t; thus, one can only speak of the "low" or "high" yielding district-year combinations, instead of low (high) yielding "districts" or "years" separately. However, for substantial parts of India, yields tend to have relatively low variability and only regions with low yield on average tend to have high variability. Therefore, we loosely interpret district-year combinations with low and high yields as "districts" with low and high yield. We note that such terminology is clearly inappropriate for the quantiles 0.5 and 0.75, since these can correspond to "low" ("high") years in high (low) yielding districts, and therefore, we can only refer to them by the less useful "district-year" terminology.

	rice					wheat				
	median	**CI**	**min**	**max**	**mean**	**median**	**CI**	**min**	**max**	**mean**
Scenario 1										
tau= 0.25	8.400	[8.34,8.48]	-10.860	11.548	5.638	-4.900	[-4.96,-4.83]	-11.500	1.635	-5.032
tau= 0.50	3.990	[3.95,4.04]	-11.730	5.601	2.107	-3.100	[-3.18,-3.04]	-8.911	2.141	-3.298
tau= 0.75	1.180	[1.16,1.21]	-10.975	2.216	-0.264	-2.240	[-2.3,-2.19]	-6.261	2.220	-2.144
tau= 0.90	0.662	[0.632,0.696]	-12.217	1.705	-0.707	-2.770	[-2.8,-2.74]	-8.614	1.275	-2.814
Scenario 2										
tau= 0.25	17.300	[17.2,17.5]	-20.280	22.193	12.603	-10.900	[-11.1,-10.7]	-22.470	3.196	-10.800
tau= 0.50	7.400	[7.32,7.52]	-22.368	10.648	4.201	-7.160	[-7.32,-7.05]	-17.790	4.212	-7.194
tau= 0.75	1.840	[1.79.1.92]	-20.649	3.949	-0.622	-5.200	[-5.31,-5.1]	-12.440	4.376	-4.675
tau= 0.90	0.347	[0.272,0.433]	-23.878	2.962	-1.987	-5.770	[-5.82,-5.73]	-16.470	2.516	-5.750

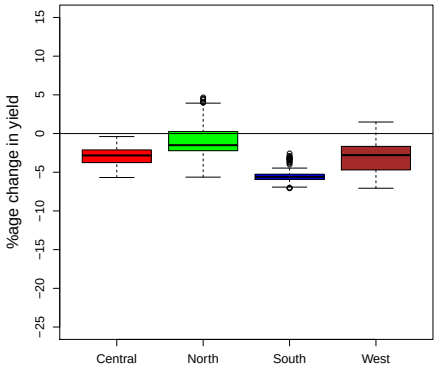
Table 2.7: Summary statistics of percentage change in yield of rice and wheat under different scenarios at the district level. For rice, scenarios 1 and 2 correspond to uniform increase in daily temperatures over the growing season by 1° and $2°C$, along with a 10% increase in seasonal rainfall, while for wheat, they correspond to increases in daily temperatures of 0.5 and 1^0C respectively. Median impacts are reported along with (time) bootstrapped confidence intervals (in []).

production, also lose significantly, up to 6% of yield with a uniform increase of $0.5°C$ and 9%, with an increase of $1°C$. Losses in the Northern region are smaller, at about $2-5\%$, whose impacts, however, can translate into large welfare losses due to the very high production losses embodied.

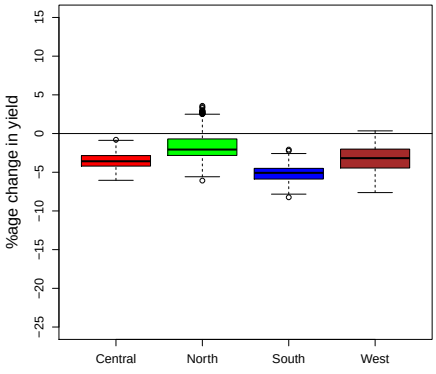


(a) 25^{th} percentile

(b) Median



(c) 75^{th} percentile

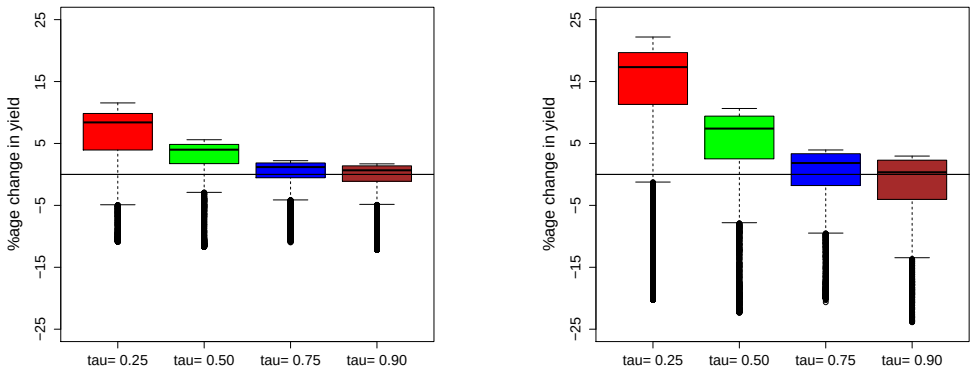


(d) 90^{th} percentile

Figure 2.4: Box plots of changes in wheat yield for different regions under scenario 1 for different quantiles.

2.5.5.2 Rice

The impact of climate change, as simulated here, on rice appears, at a first glance, from Table 2.7 to be diametrically opposite, with modest to substantial increases in yield, of $2-18\%$, except at the higher quantiles. However, two features are worth noting: median changes are substantial only at the first two quartiles, under both scenarios, and changes at the highest quantile are close to zero or even modestly negative. However, as clearly seen in Figure 2.5 (unlike in the case of wheat wherein there are only a few districts whose impacts differ in sign) a substantial number of district-years, at most quantiles, experience significant ($>5\%$) losses, indicating significant heterogeneity in impacts.



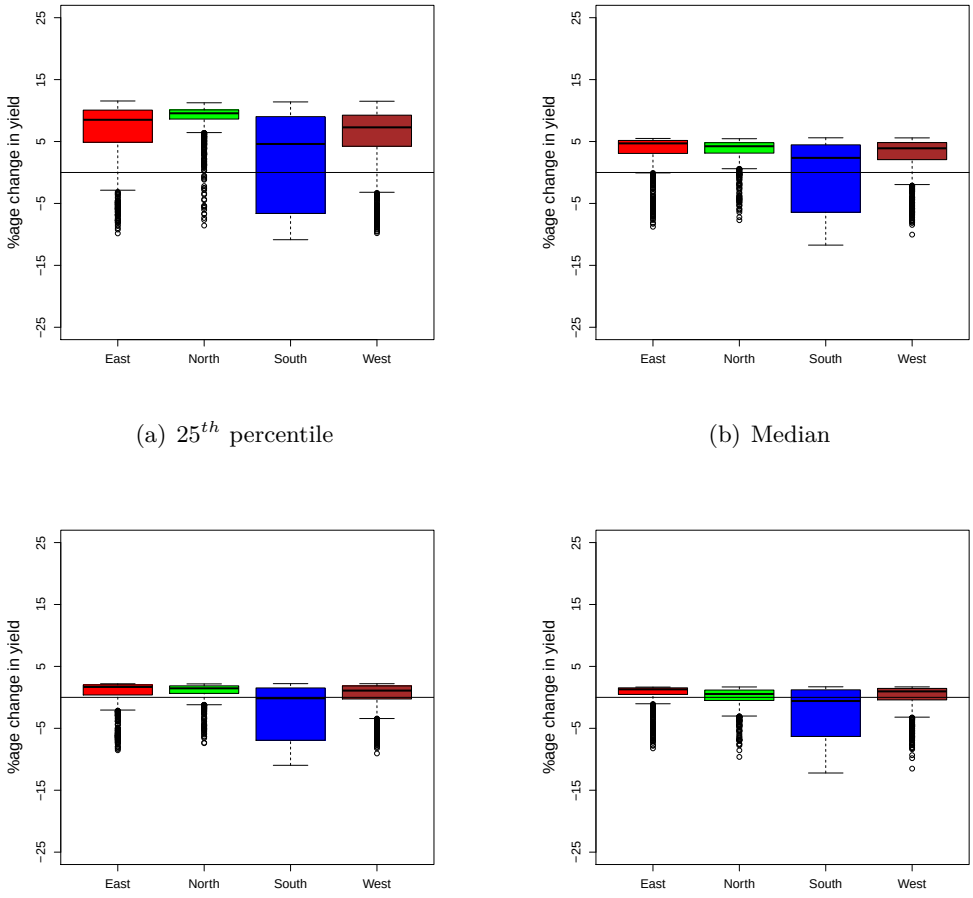
(a) Scenario 1a

(b) Scenario 1b

Figure 2.5: Box plots of changes in rice yield under scenarios indicated, for various quantiles. Scenarios are same as in Table (3)

Turning now to a closer analysis of the regional distribution of losses and gains, it is striking, from Figure 2.6 that most of the gains are driven by the Eastern region, which has a very large share in area and production of rice, while most of the losses are driven by losses in parts of the Southern region, and this is consistent across quantiles and scenarios (figures for scenario 2 not shown but are available upon request). The results of the regional

analysis help inform the pathways to possible gains. The impact of climate change on rice cultivation in Eastern India has been investigated, with some studies indicating possible increases in yield. Further, the study by Welch *et al.* [2010] already cited reports a similar increase in South-Eastern India. This study therefore adds to the weight of evidence on this important issue.



(a) 25^{th} percentile

(b) Median

(c) 75^{th} percentile

(d) 90^{th} percentile

Figure 2.6: Box plots of changes in rice yield for different regions under scenario 1 for different quantiles.

There is substantial prior literature, for India, employing crop simulation models for an

investigation of impacts of climate change on yields of various crops (see for instance Attri and Rathore [2003]; Chatterjee [1998]; Kumar and Parikh [2001]; Mall and Aggarwal [2002]; Mall *et al.* [2006a,b]). The results of such experiments have been varied, but for the crops considered here, quite clear; most studies report *minimal direct impact* (defined as simply an increase in daily temperatures) of climate change on yields of major crops, including rice, grown in the kharif season, with modest or substantial positive impacts on rice yields, in particular, for various regions (specifically, minimal impact on Northwestern and Central India). This is consistent with the results obtained here.

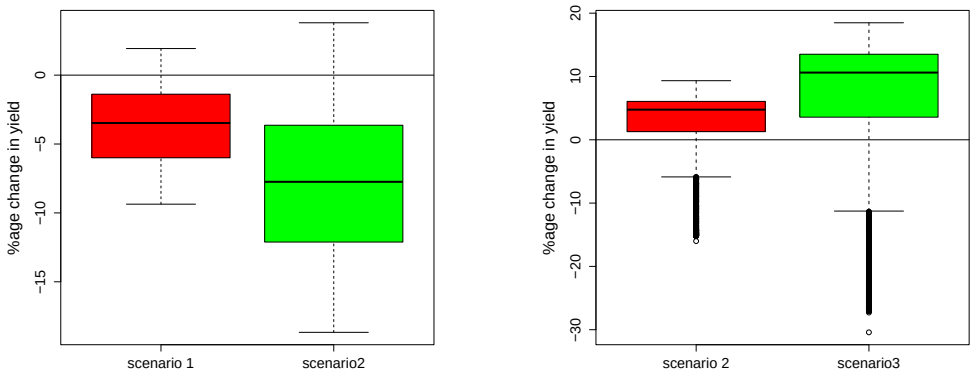
Second, most studies report significant *negative* impacts on crops grown in the *rabi* (winter) season, primarily due to increase in night time temperature. Our results also indicate, for the moderate scenarios considered here, substantially negative impacts on wheat yields, mostly in the Southern and Central regions, which is consistent with the results obtained in many crop modeling efforts, and modest increases in rice yield, driven for the most part by increases in the Eastern and Northern regions.

2.5.6 Comparison with the mean (fixed effects panel) regression

In order to compare the predictions obtained from a FE-QR framework with those of a linear FE framework, we consider box plots of predicted changes (as for the FE-QR case) in Figure 2.7[32]. For wheat, the predicted median impact under the two scenarios, of -3 and -7% are comparable to those of the median, in Figure 2.3. However, it is evident that under the reasonable scenarios considered here, mean impacts cannot fully characterize the substantial differences in impacts of changes in weather on yield. A similar result is seen with rice, from Figure 2.5, i.e. that the mean regression framework is not suited to characterizing the heterogeneity of impacts of (changes in) weather on yield. The inadequacy of FE panel

[32]In the interests of brevity, a table similar to the FE-QR case is not presented. It is, however, available upon request

regression in capturing heterogeneity in many interesting applications has been repeatedly emphasized in Galvao [2009], Harding and Lamarche [2009] and Powell [2009] and this study adds to the weight of evidence.



(a) wheat (b) rice

Figure 2.7: Box plots of changes in rice and wheat yield under scenarios indicated, from the linear fixed effects panel data model.

2.6 Conclusions

This paper employed a newly developed panel data quantile regression methodology to first estimate the relationship between current weather and agricultural yields of the two most important food grains for India, rice and wheat. The framework used allows a fuller exploration of the conditional distribution of yield beyond just the conditional mean, which is the focus of most studies. Results of the estimation, when projected on to moderate, and almost universally accepted uniform climate change scenarios for India, indicate a *significant negative* impact on wheat yields, of up to 11%, primarily in Southern and Central India, with more moderate losses in the more important Northern region. Further, these impacts were seen to be most negative on the *most productive* districts, indicating losses in production which are likely significant.

For rice, however, impacts of uniform warming of up to 2^0C were shown to be mildly positive, with increases in yield (up to 17% at the lowest quantile) concentrated in the lower quantiles, or in the least productive areas, while impacts at the upper quantile are seen to be minimal or even mildly *negative*. This is likely to translate into a moderate increase in production of rice. The results here are consistent with many estimates of changes in rice yields obtained using various crop models, under a variety of climate change scenarios, as well as a recent study of rice cultivation in Asia under irrigated conditions (Welch *et al.* [2010]).

These results indicate significant losses in wheat production, and possibly moderate gains in rice production. However, given the substantial number of outliers in rice with yield losses, these estimates likely translate into potentially substantial decreases in rice production, at least at the local scale. In summary, this study indicates that in the absence of significant changes in agricultural practices and technology, climate change is likely to lead to increases in food insecurity for India's poor, as a result of decreased yield at the national (or regional) level.

Bibliography

S. Attri and L. Rathore. Simulation of impact of projected climate change on wheat in india. *International Journal of Climatology*, (23):693–705, 2003.

M. Buchinsky. Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study* 1. *Journal of Econometrics*, 68(2):303–338, 1995.

A. Chatterjee. Simulating the impact of increase in termperature and co2 on growth and yield of maize and sorghum. *Unpublished MSc Thesis*, 1998.

R. Davidson and J.G. MacKinnon. Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis*, 51(7):3259–3281, 2007.

O. Deschenes and M. Greenstone. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, 97(1):354–385, 2007.

E. Duflo and R. Pande. Dams. *The Quarterly Journal of Economics*, 122(2):601–646, 2007.

J. Fasullo and PJ Webster. A hydrological definition of Indian monsoon onset and withdrawal. *Journal of Climate*, 16:3200–3211, 2003.

A. Fisher, M. Hanemann, M.J. Roberts, and W. Schlenker. Climate Change and Agriculture Reconsidered. Technical report, mimeo, 2009.

B. Fitzenberger, R. Koenker, and J.A.F. Machado. *Economic applications of quantile regression*. Physica Verlag, 2002.

B. Fitzenberger. The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics*, 82(2):235–287, 1998.

A. Galvao. Quantile regression for dynamic panel data with fixed effects. Technical report, mimeo, University, 2009.

S. Gonçalves and H. White. Bootstrap standard error estimates for linear regression. *Journal of the American Statistical Association*, 100(471):970–979, 2005.

S. Gonçalves. The moving blocks bootstrap for panel linear regression models with individual fixed effects. Technical report, Working Paper, Department of Economics, University of Montreal, 2008.

R. Guiteras. The impact of climate change on indian agriculture. *Unpublished manuscript*, 2008.

M. Harding and C. Lamarche. A quantile regression approach for estimating panel data models using instrumental variables. *Economics Letters*, 104(3):133–135, 2009.

A.V.M. Ines, J.W. Hansen, and A.W. Robertson. Enhancing the utility of daily GCM rainfall for crop yield prediction. *International Journal of Climatology*, 2010.

G. Kapetanios. A bootstrap procedure for panel data sets with many cross-sectional units. *Econometrics Journal*, 11(2):377–395, 2008.

K. Kato, A.F. Galvao Jr, and G.V. Montes-Rojas. Asymptotics and bootstrap inference for panel quantile regression models with fixed effects. 2010.

R. Koenker, P. Ng, and S. Portnoy. Quantile smoothing splines. *Biometrika*, 81(4):673, 1994.

R. Koenker. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89, 2004.

R. Koenker. *Quantile regression*. Cambridge Univ Pr, 2005.

R. Koenker. Additive Models for Quantile Regression: An Analysis of Risk Factors for Malnutrition in India. *Advances in Social Science Research Using R*, pages 23–33, 2010.

K. Kavi Kumar and Jyoti Parikh. Indian agriculture and climate sensitivity. *Global Environmental Change*, 1:147–154, 2001.

C. Lamarche. Robust penalized quantile regression estimation for panel data. *Journal of Econometrics*, 2010.

T. Lancaster. The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2):391–413, 2000.

RK Mall and PK Aggarwal. Climate change and rice yields in diverse agro-environments of India. I. Evaluation of impact assessment models. *Climatic Change*, 52(3):315–330, 2002.

R. K. Mall, Ranjeet Singh, Akhilesh Gupta, G. Srinivasan, and L. S. Rathore. Impact of climate change on indian agriculture. *Climatic Change*, (78):445–748, 2006.

RK Mall, R. Singh, A. Gupta, G. Srinivasan, and LS Rathore. Impact of climate change on Indian agriculture: A review. *Climatic Change*, 78(2):445–478, 2006.

R. Mendelsohn, W.D. Nordhaus, and D. Shaw. The impact of global warming on agriculture: a Ricardian analysis. *The American Economic Review*, 84(4):753–771, 1994.

Vincent Moron and Andrew. W Robertson. Interannual variability of summer monsoon rainfall onset date across India. *Submitted to Geophysical Research Letters*, 2009.

David Powell. Unconditional Quantile Regression for Panel Data with Exogenous or Endogenous Regressors. 2009.

M. Rajeevan, J. Bhate, JD Kale, and B. Lal. Development of a high resolution daily gridded rainfall data for the Indian region. *Met. Monograph Climatology*, 22:2005, 2005.

A. Rosen. Set identification via quantile restrictions in short panels. 2009.

M.R. Rosenzweig and H.P. Binswanger. Wealth, weather risk and the composition and profitability of agricultural investments. *The economic journal*, 103(416):56–78, 1993.

A. Sanghi, R. Mendelsohn, and A. Dinar. Measuring the impact of climate change on indian agriculture. Technical Report 402, World Bank, 1998.

W. Schlenker and M.J. Roberts. Nonlinear effects of weather on corn yields. *Applied Economic Perspectives and Policy*, 28(3):391, 2006.

W. Schlenker and M.J. Roberts. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594, 2009.

W. Schlenker, W.M. Hanemann, and A.C. Fisher. The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1):113–125, 2006.

AK Srivastava, M. Rajeevan, and SR Kshirsagar. Development of a high resolution daily gridded temperature data set (1969-2005) for the Indian region. *Atmospheric Science Letters*, 10(4):249–254, 2009.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):pp. 267–288, 1996.

J.R. Welch, J.R. Vincent, M. Auffhammer, P.F. Moya, A. Dobermann, and D. Dawe. Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures. *Proceedings of the National Academy of Sciences*, 107(33):14562, 2010.

Appendix 2.A Precipitation measures

The monsoon is, at its most basic, a large scale atmospheric phenomenon which determines moisture availability and whose interaction with local features determines actual local rainfall. Therefore, a definition of "onset" of the monsoon is only climatologically meaningful at a suitable large scale, generally much larger than a district. This is the major reason that the IMD defines onset as occurring at a particular location on the Southwestern coast of India (Moron and Robertson [2009] and references therein). On the other hand, these definitions of the monsoon are completely divorced from local rainfall and are uninformative to the farmer.

Recently, there has been renewed interest among the Climatologists in attempting to hydrologically define onset of the monsoon for India. Fasullo and Webster [2003] provide a definition of onset in terms of vertically integrated moisture transport, which is valid only for larger regions than considered here. Moron and Robertson [2009] provide a more local, at the grid-level, definition of monsoon onset and provide evidence that it does correspond to actual onset at selected locations. Their measure of onset, at the grid-level, maybe defined simply as the first day after April 1 that rainfall for X days exceeds a certain threshold without being followed by Y days in succession without rainfall. The rationale for this is straightforward: onset is a phenomenon involving rainfall on a few successive days, while the second condition ensures that there are no "false starts". As before, we compute district-level onset from grid-level onset. Finally, in order to proxy for substantial periods with no rainfall, we compute the number of "dry spells" in the season, wherein a dryspell is defined as at least 7 days with rainfall below $5mm$, a very low threshold. Again, this is computed at the district level.

Appendix 2.B Details of the Bootstrap Procedure used

For unbalanced panels, which in practice are the commonest forms of panel data, the modifications to the time bootstrap are straightforward but algebraically tedious and will not be illustrated here. We will however point out that the same method outlined above maybe applied with resampling now being specific to each individual unit. In the absence of dependence in the error structure, we may simply pursue a time bootstrap. However, given (at least in our application), the suspicion of strong dependence in both space and time, such time-bootstrap-based inference will clearly be biased (in uncertain directions). A general but heuristic solution, when spatial dependence is suspected, is to use the time bootstrap and when spatial dependence is suspected, to use the cross-section bootstrap (as indicated in for instance Kapetanios [2008]). In the fixed effects quantile regression setting it is not clear that there is any asymptotic justification for this approach to the covariance matrix estimation (due to the lack of theoretical results on the second order accuracy of the bootstrap).

Denoting the individual coefficients (on the covariates of interest) β,which is of dimension $K \times 1$, the estimated variance of the k^{th} component of β as $\hat{\sigma}^2_{ii}$, $n = \sum_i^N T_i$, the sample size, we have that the $t - statistic$ for the test $H_0 : \beta_k = 0$ of the k^{th}coefficient is:

$$t_k = \frac{\widehat{\beta_k}}{\widehat{se_k}}$$

$\widehat{se_k} = \sqrt{\frac{\overline{C^*_{n,k}}}{n}}$where, in the notation of Gonçalves and White, 2005, $\overline{C^*_{n,k}}$ is the $(k,k)^{th}$ element of the bootstrapped variance matrix $\overline{C^*_n}$with

$$\overline{C^*_n} = \frac{n}{B}\sum_{j=1}^{B}\left(\beta^{\hat{(j)}} - \beta^a\right)\left(\beta^{\hat{(j)}} - \beta^a\right)^T \tag{2.13}$$

where the pivotal values used, β^a, is the bootstrapped mean i.e.

$$\beta^a = \overline{\beta^*} = \frac{1}{B}\sum_{j}^{B} \hat{\beta^{(j)}}$$

A final point to note is that one may obtain the critical values by bootstrapping the variance of $\sqrt{n}\beta_k^*$, an approach corresponding to the so-called "double bootstrap". A major drawback of this approach is the computational burden imposed, due to which we do not pursue the approach in the present instance. The simplified "Fast Double Bootstrap" of Davidson and MacKinnon [2007] is likely to provide, in this case, the required improvements without the same computational burden and we leave this approach to future research.

Note again that the validity of the above approach ,including the double bootstrap, has been analyzed in detail in Gonçalves and White, 2005 for the case of the cross-section regression models, and they conjecture that the results are likely to be valid for the case of cross-sectional quantile regression. If such is the case, then it is likely to also be valid for the case of quantile regression for fixed-effects panel data, under slightly different conditions.

Appendix 2.C Supplementary Tables

	South			East			West			North		
	median	**min**	**max**	**median**	**min**	**max**	**median**	**min**	**max**	**median**	**min**	**max**
Scenario 1												
tau= 0.25	4.600	-10.860	11.397	8.510	-9.838	11.548	7.290	-9.806	11.506	9.560	-8.571	11.257
	[4.35 , 5]			[8.31 , 8.72]			[7.17 , 7.43]			[9.51 , 9.61]		
tau= 0.5	2.360	-11.730	5.601	4.700	-8.767	5.495	3.900	-10.035	5.588	4.240	-7.707	5.453
	[2.12 , 2.54]			[4.66 , 4.75]			[3.81 , 4]			[4.2 , 4.29]		
tau= 0.75	-0.148	-10.975	2.216	1.680	-8.545	2.195	1.090	-9.059	2.207	1.460	-7.402	2.174
	[-0.349 , -0.0366]			[1.63 , 1.73]			[1.02 , 1.17]			[1.42 , 1.51]		
tau= 0.9	-0.590	-12.217	1.705	1.290	-8.236	1.659	0.965	-11.510	1.694	0.545	-9.622	1.681
	[-0.698 , -0.49]			[1.26 , 1.32]			[0.914 , 1.03]			[0.498 , 0.597]		
Scenario 2												
tau= 0.25	12.700	-20.280	22.186	19.400	-15.838	22.145	16.600	-15.477	22.193	18.300	-12.396	21.636
	[12.2 , 13.2]			[19.2 , 19.5]			[16.2 , 16.8]			[18.2 , 18.5]		
tau= 0.5	4.400	-22.368	10.648	9.220	-13.559	10.407	8.410	-19.801	10.637	6.560	-17.596	10.470
	[4.13 , 4.7]			[9.09 , 9.33]			[8.25 , 8.53]			[6.41 , 6.7]		
tau= 0.75	-0.715	-20.649	3.941	3.110	-14.240	3.890	2.610	-17.510	3.949	1.360	-16.044	3.903
	[-0.953 , -0.42]			[3.01 , 3.17]			[2.53 , 2.74]			[1.25 , 1.47]		
tau= 0.9	-1.800	-23.878	2.932	1.710	-13.330	2.962	1.810	-21.750	2.941	-1.160	-20.407	2.908
	[-2.17 , -1.4]			[1.57 , 1.8]			[1.69 , 1.92]			[-1.31 , -0.993]		

Table 2.8: Summary statistics of percentage change in yield of rice, for different regions, under different scenarios, at the district level. Scenario 1 and 2 correspond to uniform increase in daily temperatures over the growing season by 1 and 2^0C, along with a 10% increase in seasonal rainfall respectively. Median impacts are reported along with (time) bootstrapped confidence intervals (in []).

	South			North			Central			West		
	median	**min**	**max**	**median**	**min**	**max**	**median**	**min**	**max**	**median**	**min**	**max**
Scenario 1												
tau= 0.25	-10.6	-11.5	-0.8788	-3.16	-11.25	1.633	-5.99	-11.23	1.635	-6.26	-9.956	1.635
	[-10.6 , -10.6]			[-3.29 , -3.06]			[-6.14 , -5.91]			[-6.49 , -5.98]		
tau= 0.5	-8.1	-8.89	0.3793	-1.59	-8.911	2.141	-4.1	-8.859	2.141	-4.3	-7.578	2.135
	[-8.12 , -8.08]			[-1.68 , -1.5]			[-4.24 , -4.03]			[-4.52 , -4.04]		
tau= 0.75	-5.12	-5.959	0.8055	-0.964	-6.261	2.21	-3.15	-6.254	2.22	-3.19	-5.421	2.21
	[-5.2 , -5.04]			[-1.04 , -0.882]			[-3.24 , -3.06]			[-3.36 , -2.93]		
tau= 0.9	-5.43	-7.099	-0.5471	-1.89	-8.225	1.267	-3.68	-8.614	1.275	-3.22	-5.078	1.268
	[-5.52 , -5.34]			[-1.93 , -1.83]			[-3.75 , -3.61]			[-3.35 , -3.1]		
Scenario 2												
tau= 0.25	-21.1	-21.99	-2.788	-7.44	-22.47	3.192	-13	-22.37	3.196	-13.5	-20.22	3.178
	[-21.1 , -21.1]			[-7.7 , -7.3]			[-13.3 , -12.9]			[-13.8 , -12.9]		
tau= 0.5	-16	-17.12	-0.09675	-4.16	-17.79	4.212	-9.17	-17.64	4.211	-9.45	-15.36	4.203
	[-16.1 , -16]			[-4.34 , -3.96]			[-9.43 , -8.92]			[-9.82 , -9]		
tau= 0.75	-9.63	-11.74	0.8845	-2.72	-12.44	4.376	-7.05	-12.36	4.376	-7	-10.71	4.376
	[-9.8 , -9.49]			[-2.87 , -2.5]			[-7.16 , -6.83]			[-7.3 , -6.62]		
tau= 0.9	-10.5	-14.04	-1.435	-4.07	-16.47	2.489	-7.55	-16.23	2.516	-6.51	-10.08	2.488
	[-10.7 , -10.3]			[-4.22 , -3.92]			[-7.74 , -7.45]			[-6.74 , -6.29]		

Table 2.9: Summary statistics of percentage change in yield of wheat, for different regions of India, under different scenarios, at the district level. Scenario 1 and 2 correspond to uniform increase in daily temperatures over the growing season by 0.5 and 1C respectively. Median impacts are reported along with (time) bootstrapped confidence intervals (in []).

Table 2.10: Quantile regression with fixed-effects for wheat, using Avg. Temp. instead of GDD. All regressions include region-specific cubic time trends and district fixed-effects whose coefficients are not reported. T-statistics for the test $H_0 : \beta = 0$ based on the asymptotic variance matrix are reported.

	τ=0.25	τ=0.5	τ=0.75	τ=0.9
Avg Temp	5.35E-01	5.00E-01	3.79E-01	3.73E-01
	[4.328]	[5.467]	[4.858]	[4.12]
Avg Temp sqr	-1.51E-02	-1.34E-02	-1.02E-02	-1.06E-02
	[-5.039]	[-6.077]	[-5.211]	[-4.443]
seasonal rain	5.41E-04	2.82E-04	6.49E-04	3.52E-04
	[2.478]	[1.841]	[3.745]	[1.61]
rain sqr	-1.36E-06	-9.77E-07	-1.78E-06	-2.10E-07
	[-2.032]	[-2.461]	[-3.248]	[-0.2774]
hgdd	-6.14E-05	-1.79E-04	3.92E-04	1.73E-03
	[-0.06638]	[-0.2313]	[0.5338]	[1.811]
irri. ratio	1.01E+00	9.81E-01	9.06E-01	7.76E-01
	[15.04]	[15.59]	[12.45]	[8.669]

Table 2.11: Quantile regression with fixed-effects for rice, using Avg. Temp. Other conventions same as for Table 2.10.

	τ=0.25	τ=0.5	τ=0.75	τ=0.9
Avg Temp	-7.11E-01	-5.94E-01	-4.28E-01	-4.08E-01
	[-5.167]	[-4.248]	[-3.177]	[-2.507]
Avg Temp sqr	1.33E-02	1.08E-02	7.49E-03	7.10E-03
	[5.609]	[4.439]	[3.141]	[2.446]
seasonal rain	-8.43E-05	-3.36E-05	3.43E-06	-3.19E-06
	[-1.798]	[-0.8897]	[0.09497]	[-0.08149]
rain sqr	1.59E-08	3.99E-09	-3.91E-09	1.42E-09
	[1.476]	[0.475]	[-0.4787]	[0.1289]
onset day	-2.58E-04	-2.09E-04	-4.91E-04	-4.09E-04
	[-1.036]	[-1.036]	[-2.557]	[-2.074]
dry spells	-5.76E-03	-1.79E-03	-3.37E-03	-4.07E-03
	[-1.364]	[-0.5698]	[-1.13]	[-1.323]

chapter 3

Analysis of changes in extremes of rainfall over India[1]

[1] Joint work with Upmanu Lall, published as: Krishnamurthy, Chandra Kiran B., Upmanu Lall, Hyun-Han Kwon, 2009: Changing Frequency and Intensity of Rainfall Extremes over India from 1951 to 2003. *Journal of Climate*, 22, 4737 – 4746.

3.1 Introduction

Anthropogenic climate change poses potentially significant risks for the Indian subcontinent through changes in extreme rainfall characteristics. General Circulation Models (GCMs) of climate have had only limited success in reproducing the key attributes of the intra-seasonal and inter-annual variations in the Indian monsoon. Consequently, it is not clear whether GCM simulations forced with the Intergovernmental Panel on Climate Change(IPCC)-style anthropogenic change scenarios adequately represent changes in Indian rainfall extremes, especially for extreme rainfall that translates into floods or for multi-day dry periods that impact crop yield. Recently, a few papers (Guhathakurta and Rajeevan [2008],Goswami *et al.* [2006])have investigated the trends in selected extreme rainfall attributes from a daily rainfall data set that has become available through the Indian Meteorological Department (IMD). Such analyses provide a useful backdrop for assessing whether forced GCM simulations, such as those by May [2004], Kumar *et al.* [2006], among others, provide plausible scenarios for changes in extreme rainfall in the 21st century.

This paper presents an exploratory, spatially distributed analysis of the nature of monotonic trends in selected statistics of daily rainfall across India. The research presented differs from recent work on the issue(Goswami *et al.* [2006], Guhathakurta and Rajeevan [2008],Joshi and Rajeevan [2006],Rajeevan *et al.* [2008],Alexander,L.V., et al [2006], Klein Tank, A.M.G., et al [2006],Kumar *et al.* [2006] and May [2004]) in the specific statistics (frequency and intensity) of extremes considered;in the use of a non-parametric monotonic trend analysis (instead of a linear trend analysis, which is non-robust to outliers, a concern in analyzing data on extremes), and in analyzing the complete spatially distributed data set instead of an aggregate region. Further, we use methods for field significance analysis of spatial trends in each statistic and also document the concordance of trends across variables. Trends in the frequency with which daily rainfall exceeds selected thresholds, as well

as in the intensity (magnitude of such rainfall events) are considered at each grid cell over India. The rainfall amount considered to define the exceedance events corresponds to fixed percentiles of the long term rainfall data at that grid cell, and hence the threshold magnitude varies from grid cell to grid cell. Thus, changes in the local climatology of extremes are explored, rather than the rate of occurrence of a fixed extreme magnitude across a region or even the entire country.

Further, analyzing trends in frequency and intensity separately is of interest since it is possible that the number of extreme events could increase without a corresponding increase in the intensity of each event (Trenberth [1999]) and each measure provides information regarding different aspects of extreme rainfall. For instance,rainfall-indexed insurance is being introduced by several organizations in India (Gine *et al.* [2007]) and the determination of a fair premium, and the associated payout structure, requires an assessment of whether the upper tail of the probability distribution of daily rainfall is changing at the specific location where contracts are likely to be written. The need to inform these and similar applications motivates our spatially distributed analysis of trends in the exceedance of specific percentiles of the local distribution of daily rainfall.

In the monsoonal setting (Indian or Asian Monsoon), there have been a few studies focused exclusively on trends in extreme rainfall and most of these have been based on Greenhouse Gas forced model-based scenarios of the IPCC for the 21st century (Lal *et al.* [2000], Bhaskaran and Mitchell [1998], May [2004], Kumar *et al.* [2006]). Keeping in mind the biases in the models (indicated in Kumar *et al.* [2006]), we note that most models appear to predict enhanced summer monsoonal precipitation over parts of north-western India, while predicting little or no change over much of Peninsular India (Kumar *et al.* [2006]). Climate model-based studies appear to indicate an increase in the geographic extent of intense events but not necessarily an intensification of extreme events in areas already subject to high rainfall (which tend to be along the south western coast or the north-eastern

sub-himalayan region(Kumar *et al.* [2006]). Model results also indicate intensification of rainfall in most of India except parts of central and north-eastern India (May [2004]), with the most intense (maximum 24-hr rainfall) rainfall events predicted to occur over northeastern and northwestern India.

The literature in this context is typically focused on an analysis of the intensity of extreme precipitation events, such as the 1-day event or the 5-day event. The frequency of events exceeding a certain threshold has not usually been a focus. Exceptions are May [2004], in which an analysis of the frequency of number of wet days, but not of wet days wherein a certain threshold is exceeded is carried out and Goswami *et al.* [2006].

The study by Goswami *et al.* [2006], using the same gridded data set as here, reports an increasing trend in the frequency of extreme precipitation events, defined as events exceeding the thresholds of 100 and 150 mm, using pooled data from all grid cells over the central Indian region (the so-called monsoon belt), and also indicates an increase in the intensity of precipitation, as measured by the raw values of the 99.5th and 99.75th percentile of the rainfall distribution, over the same region.

The study by Guhathakurta and Rajeevan [2008] attempts to evaluate trends in rainfall over India (not rainfall extremes) using a newly-constructed data set of monthly data from 1901-2003.The results of this study appear to corroborate our findings of existence of trend-in their case, in sub-divisional rainfall-and although the difference in data, objectives make comparisons difficult, it is nonetheless interesting that this study also reports no trend in all-india data but trends in both directions in the regional analysis.

Klein Tank, A.M.G., et al [2006], Alexander,L.V., et al [2006] and Joshi and Rajeevan [2006] all use (very different) station-level data sets to carry out an analysis of (various measures of) extreme events. Joshi and Rajeevan [2006] use station data (about 199 stations from 1901-2000)for India to carry out a linear, parametric trend analysis on various measures of extremes. They find increasing trends for certain regions (West coast and Northwestern

India) as well as an increase (as in Goswami *et al.* [2006]) in the contribution of heaviest rains to total rainfall. Finally, Rajeevan *et al.* [2008], using a longer station-level data set (1901-2004), carry out an analysis very similar to Goswami *et al.* [2006],over a slightly different region, and find increasing trends (after accounting for inter-decadal variations in the extreme events) in both heavy and very heavy rainfall events (as defined in Goswami *et al.* [2006]). They also make a preliminary attempt at linking such trends to ocean surface temperatures.

Klein Tank, A.M.G., et al [2006] use station data (for South and Central Asia) for two different time periods (1961-2000 and 1901-2000) to obtain linear, parametric trends at each station (as well as simple regional average trends) for 8 measures of extremes. They report no coherent trends in most measures of precipitation extremes at the regional level. Alexander,L.V., et al [2006], on the other hand, carry out a global analysis using station data for various regions of the world. Various measures of extremes are computed at the station level and are then gridded (2.5 latitude by 3.75 longitude) before an analysis very similar to that carried out here is attempted for various measures of extremes.

3.2 Data

This study utilizes a recently available gridded daily data-set for India (Rajeevan *et al.* [2006]), consisting of 1300 grid cells, each 1^0 lat x1^0 long, for 53 years (1951-2003), available from the IMD. Of these 1300 grids, 357 grids covering all of India's land area were used for the analysis. This is the same data set from which Goswami *et al.* [2006] draw their subset for analysis.

3.3 Definition of statistics of extremes

We consider two measures of extremes, frequency and intensity, defined respectively as the number of days with rainfall events (each year) exceeding a threshold and the average daily rainfall (for each year) on the days on which rainfall exceeds the specific threshold.

A threshold is defined in terms of a fixed percentile (two were considered, the 90th and the 99th) of the daily rainfall series at the grid cell, considering only days with non-zero rainfall. As a result, the threshold magnitude varies from grid cell to grid cell (but not year to year). A time series of frequency at each grid cell is computed as:

$$f_{jt} = \sum_{i=1}^{365} 1_{P_{itj} > P_j^*}$$

where t is the year, j the grid box, P_{itj} the rain on day i in year t and grid j, and P_j^* is the rainfall threshold for grid j, 1 is an indicator function that takes the value 1 if the argument is true and 0, else.

Correspondingly, the intensity time series is derived as (following the same notation)

$$r_{jt} = \left(\frac{\sum_{i=1}^{365} 1_{P_{itj} > P_j^*} P_{ijt}}{f_{jt}} \right) 1_{f_{jt} > 0}$$

For each grid cell, the number of non-zero precipitation events each year was identified and the 90^{th} (99^{th}) percentile of this series estimated. The median of these 90^{th} (99^{th}) percentile values across all years was then selected to be the threshold for that particular grid cell. The spatially varying climatology of extreme rainfall across India is thus addressed(see also [Joshi and Rajeevan, 2006, 6]).

We feel that this procedure better represents the spatial aspects of the monsoon process than a threshold fixed across grids, since the monsoon rainfall varies substantially across India, and we are interested in how the spatial pattern may of extreme rainfall may have

changed across the country. This is evident from figure 3.1, which illustrates that the spatial pattern of the thresholds are very similar to the monsoonal precipitaion patterns, with the largest thresholds obtained at the south-western, western coast and the north-eastern region.

The primary analyses were carried out separately for data for the monsoon season, June-July-August-September, and for the rest of the year. The monsoon season results are reported here.

3.4 Trend Analysis

The Mann-Kendall (MK) test is used for the detection of monotonic trends in the derived frequency and intensity data for each grid cell. An estimate of the Sen slope, a robust estimate of the monotonic trend, is also computed, along with its significance level. The MK test is a rank-based test, with no assumptions as to the underlying probability distribution of data ([Helsel and Hirsch, 1992, 326-327]).The test statistic, computed based on pairwise comparison between the values of a series, is asymptotically normally distributed, independent of the distribution of the original series. A robust estimate of the magnitude of the slope of the trend is estimated using the method of Sen, as the median of pairwise slopes between elements of the series([Yue *et al.*, 2002, 16-17]).

For each grid cell, and separately for the frequency and the intensity data, we test (at the 10% significance level) (a) the null hypothesis of no trend (b) the null hypothesis of no increasing trend and (c) the null hypothesis of no decreasing trend. Recognizing that a certain number of rejections of the null hypothesis are to be expected given the large number of tests conducted, we construct a field significance test (described in the next section) to assess whether the outcomes of the significance tests at the grid level may be consistent with what is expected purely by chance. Here, we examine the general features of the trends revealed by the MK test.

Figure (3.2) provides the spatial distribution of the trends for grids where the null

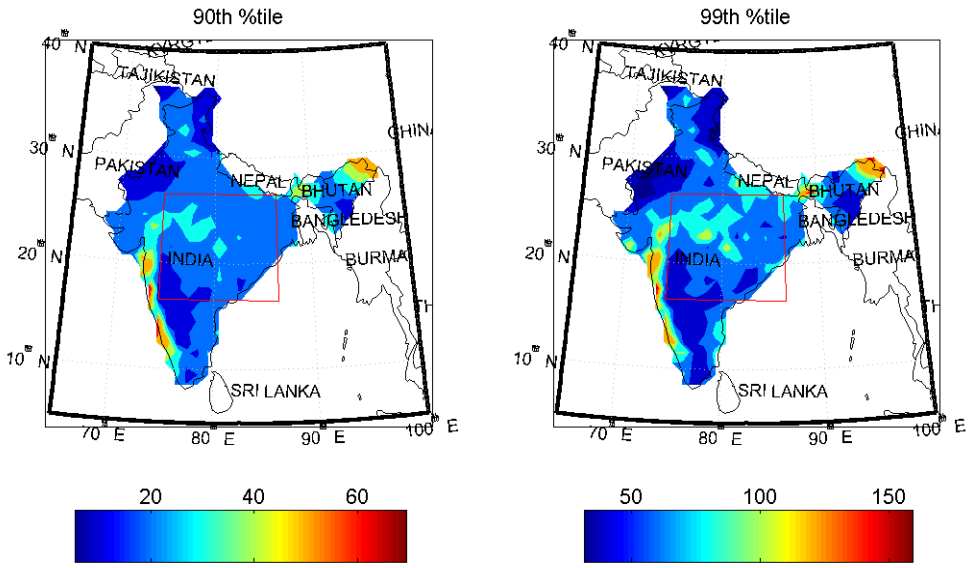


Figure 3.1: Contours of exceedance thresholds of daily rainfall (mm) at each grid. The boxed area indicates the region considered in Goswami *et al.* [2006]

hypothesis of no monotonic trend is rejected at the 10% significance level, while table 3.1 provides a tabulation of the number of such grids.

Even at the higher threshold, however, the numbers of negative trends are quite noticeable, especially in frequency. Further, it may be inferred from figure(3.1) that the spatial distribution of positive and negative trends are noticeably prominent in two particular regions (in all cases); the Eastern region and the Northern region, while the Western and Peninsular regions appear to be relatively less densely populated by significant trends. In addition, in much of the Southern region, the coasts (or the immediate interiors of the coast) appear to be the region with changes[2].

For exceedances of the 90^{th} percentile, the number of decreasing trends in frequency dominates the number of increasing trends. This observation runs counter to the assessments reported in the literature, where increasing trends in extremes are the focus. For instance, Goswami *et al.* [2006] and Kumar *et al.* [2006] find only increasing trends (in the first case over a restricted subset of the domain investigated here), with a fixed threshold of rainfall applied. Joshi and Rajeevan [2006], using thresholds varying with station, is the only study to report decreasing trends (at a few stations).

Note from table 3.1 that the number of increasing trends in intensity is higher than decreasing trends at the same threshold, which suggests that when exceedance of the 90^{th} percentile of grid rainfall occur, the amount of rain has been increasing, an observation likely to support the direction of trends reported in Goswami *et al.* [2006] (with a fixed rainfall threshold) and in Joshi and Rajeevan [2006] (with a spatially varying threshold).

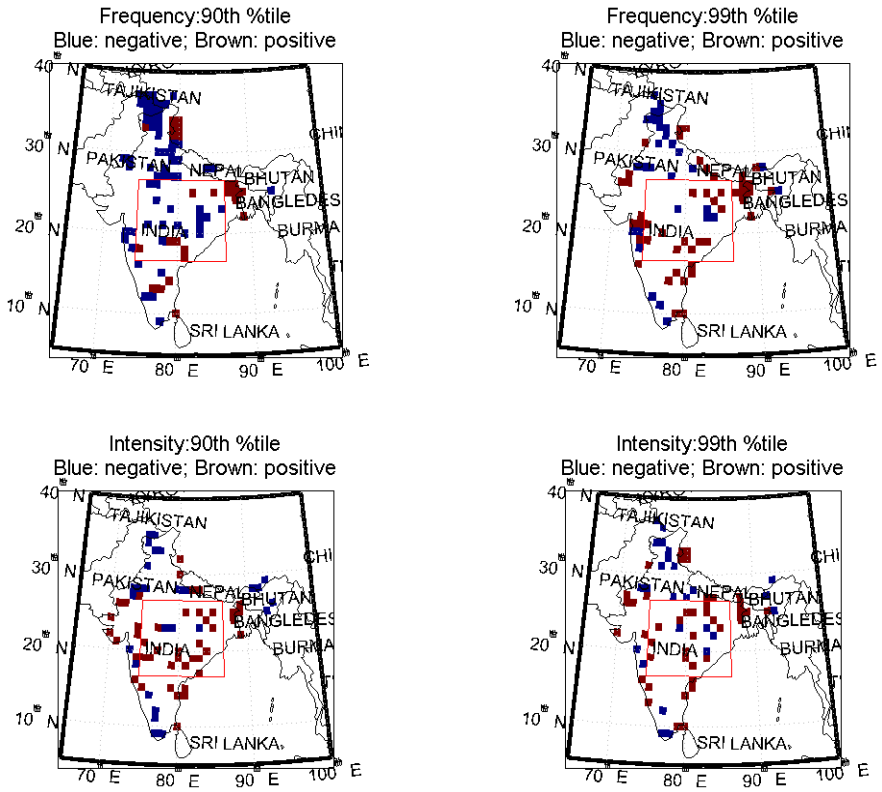


Figure 3.2: Spatial Distribution of grids where the Null of no trend is rejected by the MK test (one-sided, at the 10% significance level) (Blue indicates decreasing, Brown,increasing trends;exceedance of the 90^{th} and 99^{th} percentile of daily rainfall is considered). The boxed area indicates the region considered in Goswami *et al.* [2006]

	Number of Grids with	
	increasing trend	**decreasing trend**
90^{th}%tile threshold		
Frequency	25	61
Intensity	44	27
99^{th}%tile threshold		
Frequency	45	30
Intensity	42	25

Table 3.1: Distribution of grids with statistically significant (at the α_l=10% significance level) trends (357 grid cells in total)

3.5 Tables

A perusal of the trends in frequency and intensity of exceedance of the 99^{th} percentile threshold supports such a speculation, given the dominance of increasing trends in both frequency and intensity at this threshold. However, contrary to much of the literature, a fair number of decreasing trends are noted in our analyses. From the figures it is clear that while the details vary by threshold and metric (frequency and intensity), increasing trends dominate in the coastal regions and in the eastern region (west of Bangladesh), while decreasing trends appear to be more prevalent in the northern, central and north-eastern parts of India. Indeed, from these figures, it is difficult to argue that there has been an increase in the frequency and intensity of extreme rainfall across India.

The joint trends in frequency and intensity are investigated next. The motivation is to investigate whether, as hypothesised by Trenberth [1999] and others, trends in both frequency and intensity increase or decrease jointly. The trends in frequency and intensity (in figures (3.3) and (3.4)) that are deemed significant in the independent analyses agree completely for exceedances of the 99^{th} percentile threshold. For exceedances of the 90^{th} percentile, decreasing trends in frequency and intensity at the same location are much more

[2]Changes in the extreme northern regions, especially over the states of Jammu and Kashmir and Himachal Pradesh, are based on very scanty data and will be ignored in much of the analysis

prevalent than joint increasing trends in these two metrics. It is remarkable that at the higher threshold, there is not a single grid cell with opposite directions of trends in frequency and intensity, while at the lower threshold, grids with opposing trends are evident only in the monsoon belt. At the lower threshold, there are not very many grid cells where trends in both frequency and intensity are significant. Most grids with trends in both significant are located in the eastern part of the country.

Given the spatial structure of the Indian monsoon, it is pertinent to ask if a similar (or some) structure is evident in the trends as well. To investigate this aspect, we plot the contours of the trends calculated at each grid point, and note that if there were some spatial structure in the trends, it would be evident in the contour plots. However, a perusal of figures (3.5) and (3.6) indicates that there is no evident structure to the trends, of either sign.

Having taken a broad look at the spatial distribution and direction of trends in the frequency and intensity of extreme rainfall across India, we next examine whether the number of statistically significant trends that appear to be different from zero at the 10% significance level could occur purely by chance, in an analysis of the spatially distributed data set used here.

3.6 Field Significance Test

The question addressed in this section is whether the number of increasing or decreasing trends deemed significant at the grid-cell level analysis could occur purely by chance, taking into account the possibility that the rainfall data, and hence, the trends, have an underlying spatial structure. The answer to this question depends on the specific area or domain considered (all of India or the core monsoon region as identified in Goswami *et al.* [2006], and indicated as a box in all figures). Results for the all-India data are presented first and those for the smaller, core monsoon region are discussed next.

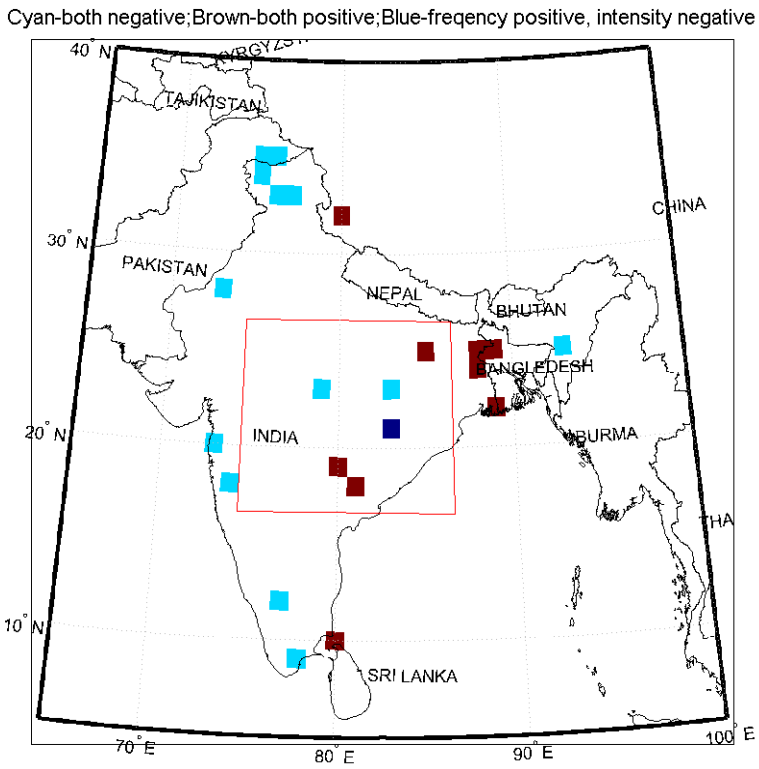


Figure 3.3: Assessment of consistency in significant trends in frequency and intensity (at the 90^{th} %tile threshold). Cyan indicates trends in both frequency and intensity decreasing, Brown indicates trends in both increasing and Blue indicates trends in frequency increasing and intensity decreasing. Only grids where trend in both frequency and intensity were statistically significant (at the 10% level) are considered. The boxed area indicates the region considered in Goswami *et al.* [2006]

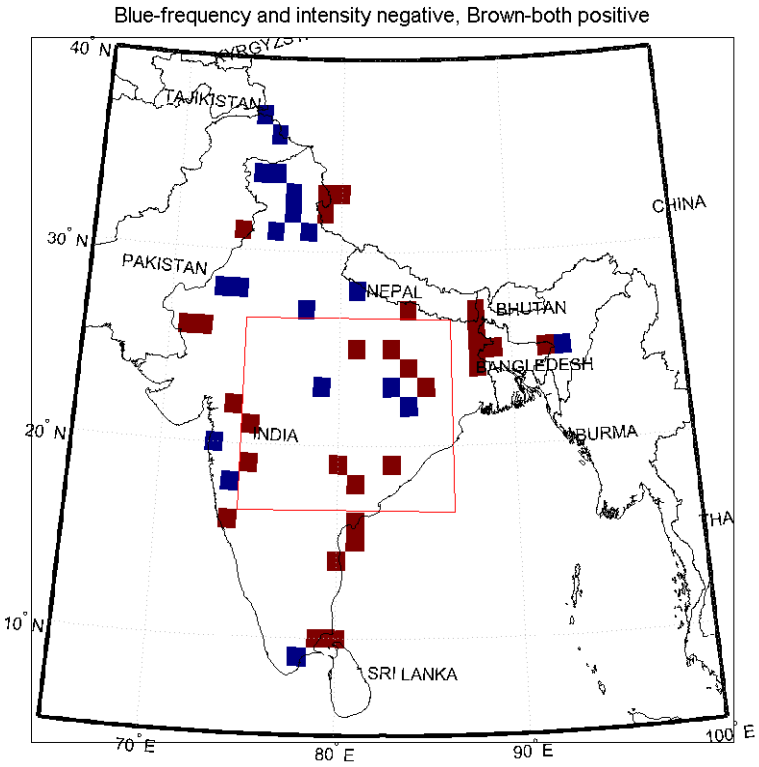


Figure 3.4: Same as figure 3 (but at the 99^{th} %tile threshold). Blue indicates trends in both decreasing, Brown indicates trends in both increasing

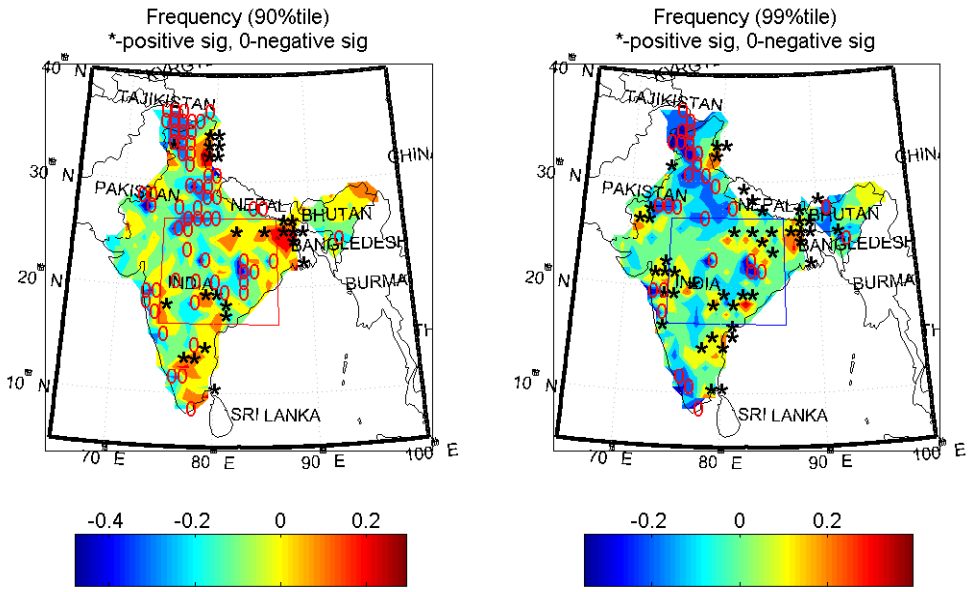


Figure 3.5: Contours of the trend estimated by the Sen Slope for frequency of exceedance at each of the grids. Grids with statistically significant trends (at the 10% level of significance) are marked with * and 0. The boxed area indicates the region considered in Goswami *et al.* [2006]

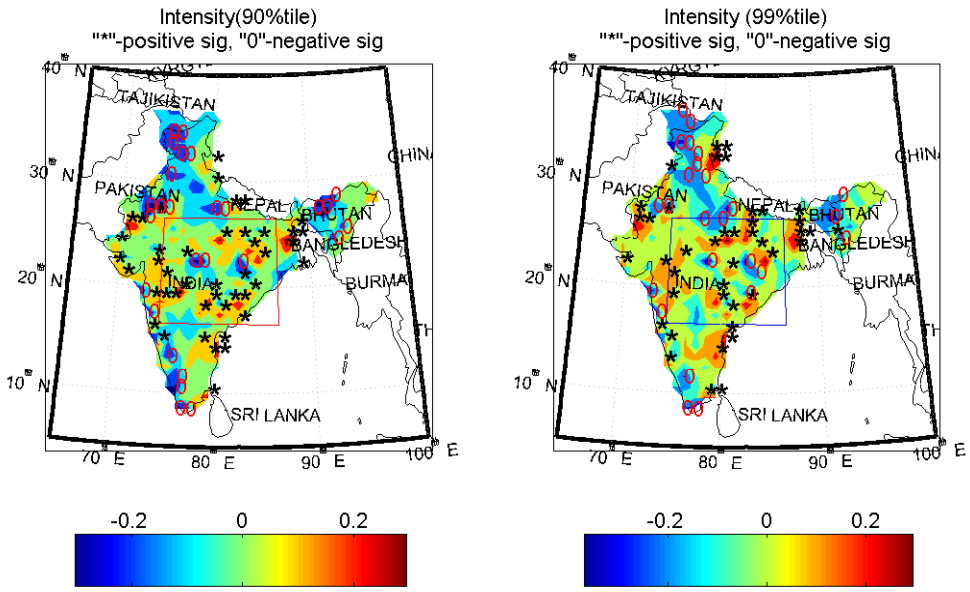


Figure 3.6: As in Fig 5 but for Intensity

The so-called field significance test (Livezey and Chen [1983]) has been typically used to address the question posed in this section. The null hypothesis of the test is that the number n out of N total grid cells exhibiting a trend at the α_l % level of significance (local or at each grid cell) is not inconsistent with the value expected by chance, considering the potential for spatial correlation across the individual time series analyzed for trends. The null hypothesis is rejected if n is larger than the number expected by chance at the α_f% significance level (global or across the domain)[3].

A one-sided test is used, to test whether the proportion of grids with significant trends, p, is greater than the (global) significance level α_f, as below

$$H_0 : p = \alpha_f \text{ v/s } H_a : p > \alpha_f$$

The test statistic

$$\frac{p - \alpha_f}{\frac{\sqrt{\alpha_f(1-\alpha_f)}}{357}}$$

is distributed $N(0, 1)$ under the null.

In most applications, the validity of the field significance test is compromised by the finiteness of the data set used and by the spatial and/or temporal correlation between the series used (Livezey and Chen [1983], Elmore *et al.* [2006], Wilks [1997]). We outline a procedure which addresses the spatial dependence of data. Spatial correlation reduces the degree-of-freedom of the test i.e. there are less than N individual realizations (of the test statistic or phenomenon; in this case, frequency and intensity) in a field of size N. Livezey and Chen [1983] estimate the effective number of realizations, $n < N$, by a Monte Carlo procedure involving generating the sampling distribution of n under the null (of field non-significance), obtaining a desired percentile (say the 5th) of this distribution, and comparing

[3]Note that the local significance level, α_l, is always taken to be 10% while the global significance level, α_f, is either 10% or 5%, depending on whether the field significance test pertains to significant trends of both (positive and negative) signs or of one particular sign

it to the minimum number of effective degree of freedom for the significance of the field, n_0. The null hypothesis is rejected if $n > n_0$.

An alternative is to generate the sampling distribution of the test statistic under the null hypothesis, keeping intact the spatial structure of the data set under consideration. The advantages of this approach include sampling the spatial correlation structure without formally specifying it (Wilks [1997]).

The bootstrap is a non-parametric method that samples, with replacement, from the original data. It is applied here by re-sampling the spatial field associated with each year, which preserves the spatial structure, but randomizes the temporal structure. The bootstrapping procedure for the field significance test was carried out by first generating 1000 random samples of 53 numbers each from 1 to 53 (with replacement). Each of these 1000 samples are then treated as realizations of a time index corresponding to 53 years of data, with the frequency and intensity spatial fields then sampled for each of these generated years.

This approach preserves the correlation structure across space but not across time. Serial correlation across years in each grid cell were not found to be significant. For each of the 1000 samples generated, the MK-test was repeated for each grid and for each sample. This leads to 1000 samples at each grid cell with a binary determination of trend significance at the α_l% level. The proportion of grid cells where significant trends at the α_l% level were found was calculated for each of the 1000 samples, to obtain 1000 realizations of the proportion of grids, p, exhibiting trends at the local significance level. This provides an estimate of the sampling distribution of the test statistic under the null hypothesis that the number of trends identified as significant at the α_l% level, across the domain, is consistent with the number expected by chance. This computation was done separately for increasing and decreasing significant trends in frequency and intensity, leading to six different tests for each exceedance threshold. The procedure and results are summarized below:

1. Mann-Kendall Trend test: Carry out the MK test, obtain the Sen slope and a count of the number of grids at which the null (of trend non-significance) was rejected (at the 10% level), compute the test statistic, denoted t_{sample}. This step is carried out for each of three types of trends (trends in both directions, increasing trends and decreasing trends only) and steps (ii) and (iii) are then repeated separately for each of these three types of trends.

2. Field significance test:

   - Randomly sample, with replacement, from the data, to obtain 1000 copies of the data matrix while retaining the spatial structure
   - Obtain 1000 realizations of the proportion of the grid cells, p, for which the hypothesis of no trend is rejected and the vector (of size 1000) of test statistics (denoted $\mathbf{t_{bootstrap}}$)
   - Construct the bootstrap estimate of the sampling distribution of the test statistic $\mathbf{T_{bootstrap}} = (\mathbf{t_{bootstrap}} - t_{sample})$. Sort this vector and obtain its $100(1-\alpha_f)^{th}$ percentile (denoted: T^*)[4]
   - test $T = t_{sample} - \alpha_f$ against T^*[5](recall that a one-sided test is employed)and reject the null hypothesis that the number of significant trends is what would be expected by chance if: $T > T^*$

3. Repeat the analysis for different thresholds

The results of the bootstrap procedure are summarized in table 3.2. First, if we consider the total number of trends (of either sign), we observe that the null hypothesis is rejected for all tests. Next, if we consider increasing trends, only in the case of frequency of exceedance

[4]This is known as the "percentile method" of bootstrap-based hypothesis testing ([Davison and Hinkley, 1997, 201-203]

[5]We note that this approach differs from the conventional one, involving testing $\mathbf{T_{bootstrap}}$ versus α; the approach adopted here is more efficient than the conventional one (Hall and Wilson [1991], Wilks [1997])

of the 90^{th}percentile is the null hypothesis not rejected. Changes in intensity indicate an increasing trend at both thresholds. The number of decreasing trends passes the significance test only for frequency at the 90^{th} percentile. Thus in summary, the hypothesis that overall there is an increasing trend in the frequency and intensity of extreme rainfall appears to have support, with the caveat that at the 90^{th} percentile, the frequency of exceedance appears to be decreasing in the central and northern regions, while the intensity of these events is increasing. A contour plot of the slopes of frequency and intensity provides a smoother representation of the nature of these trends (Figures 3.4 and 3.5).

	Frequency			**Intensity**		
	t^*	t	**Null**	t^*	t	**Null**
	90^{th} percentile threshold					
both incr. and dec.	-0.078	0.141	*reject*	-0.042	0.099	*reject*
increasing only	0.025	-0.03	do not	-0.039	0.023	*reject*
decreasing only	-0.073	0.071	*reject*	0.008	-0.024	do not
	99^{th} percentile threshold					
both incr. and dec.	-0.056	0.11	*reject*	-0.034	0.088	*reject*
increasing only	-0.039	0.026	*reject*	-0.031	0.018	*reject*
decreasing only	0	-0.016	do not	0.014	-0.03	do not

Table 3.2: Bootstrap test results (Note: (a) t: raw statistic, t^*: critical value – the $(1-\alpha_f)$quantile of the bootstrap distribution (b) α_f for increasing and decreasing trends are 0.05 while for all sig. grids, it is 0.1)(c) Null of field non-significance ($p = \alpha_f$) is rejected if $t > t^*$)

Now consider the results over the region considered by Goswami *et al.* [2006]), the main study with a similar analysis and dataset. Recall that Goswami *et al.* [2006] define extremes over a homogeneous region, use the number of days of rainfall above 100 and 150mm and the intensity of rainfall for a fixed percentile (99.5^{th} and 99.75^{th}), as measures of extremes. They report an increase in the frequency of extreme rainfall events as well as an increase in the intensity of extreme rainfall. Carrying out the field significance test outlined in the preceding section over their domain, we find that the broad conclusions from the national analysis are essentially unchanged (table 3.3 and 3.4) i.e. that while increasing trends do

exist, they are more predominant in the south-western coast and north-eastern regions, with decreasing trends being more prominent in the central regions.

	Number of Grids with	
	increasing slopes	decreasing slopes
90^{th} percentile threshold		
Frequency	7	9
Intensity	17	2
99^{th} percentile threshold		
Frequency	13	4
Intensity	12	4

Table 3.3: Distribution of grids with statistically significant (at α_l=10% significance level) trends in the region defined by Goswami *et al.* [2006] (74 grid cells in total)

	Frequency			**Intensity**		
	t^*	t	**Null**	t^*	t	**Null**
90^{th} percentile threshold						
both inc. and dec.	-0.014	0.103	*reject*	-0.068	0.143	*reject*
increasing only	0.041	-0.019	do not	-0.122	0.116	*reject*
decreasing only	0	0.022	*reject*	0.068	-0.073	do not
99^{th} percentile threshold						
both inc. and dec.	-0.041	0.116	*reject*	-0.027	0.103	*reject*
increasing only	-0.054	0.062	*reject*	-0.054	0.049	*reject*
decreasing only	0.041	-0.046	do not	0.041	-0.046	do not

Table 3.4: Bootstrap test results for slopes in region defined by Goswami *et al.* [2006] (Note: (a) t: raw statistic, t^*: critical value– the $(1-\alpha_f)$quantile of the bootstrap distribution (b) α_f for increasing and decreasing trends are 0.05 while for all sig. grids, it is 0.1)(c) Null of field non-significance ($p = \alpha_f$)is rejected if $t > t^*$

Joshi and Rajeevan [2006], using a different dataset, find increasing trends (using somewhat different measures of extremes) in very similar regions; they also report negative trends (at only two stations).

Goswami *et al.* [2006] also performed a split sample analysis, splitting the data into two parts, pre- and post-1981 and find an increasing trend in the post-1981 sample. We repeated our analysis for the same two sub-periods. The results of this analysis indicate

that the conclusions obtained using the full sample are unaltered, unlike Goswami *et al.* [2006] who find a increasing trend only in the post-1981 sample. The importance of the spatially distributed analysis performed here is that if the spatial differences noted represent non-homogeneous aspects of the monsoon, then the spatial patterns identified in the trends would potentially help inform mechanism identification and model performance evaluation.

Kumar *et al.* [2006] find significant increases in intense precipitation in much of western, northwestern and especially the south-western regions. The results of the present analysis indicates trends in mostly parts of the south western coastal regions, similar to the results of Kumar *et al.* [2006], as well as in the eastern and central regions. May [2004] reports increases over much of the Indian peninsula, while the coarse spatial resolution of the model used does not provide detail smaller over regions of India. Further, our results that decreasing trends are likely in many areas are in concurrence with May [2004] who also finds similar decreases (in the scale and/or shape parameters of the Gamma or the GPA distributions which are fit to the rainfall data) for a small number of regions. However, the spatial coarseness of the model prevents a closer spatial comparison of the results.

3.7 Discussion and Conclusion

An earlier examination of trends in extremes of Indian monsoon rainfall was developed further in this work. The analysis considered the spatial structure of changes in the extremes across the country rather than over a box (homogenous region) in central India. Broadly speaking, there is support for the hypothesis that the frequency and intensity of extreme rainfall over India may be increasing over the previous 53 years. However, there is considerable spatial variation as to the direction of change, and the spatial continuity of trends deemed statistically significant is weak. This is not unexpected, since threshold crossings are a random process, and the assessment of significance is also a threshold process. A visual examination of the spatial variation in the trends in frequency and intensity of ex-

treme rainfall suggests that the north and central sections of the Indian subcontinent have experienced a generally decreasing trend in the frequency and intensity of extremes, while the coastal regions in Peninsular India, and the region immediately west of Bangladesh have experienced increasing trends.

Even in central India, which was analyzed in aggregate by a previous study, there is some heterogeneity in the direction of the trends, and the larger scale analysis performed here helps clarify the spatial structure of the changes in the region studied in Goswami *et al.* [2006].

We do not attribute the trends observed to anthropogenic climate change or to inter-decadal climate variability which may be of natural origin. Rather, we offer the results of this analysis as a benchmark to the climate community to consider more detailed studies of the spatial structure of changes in the Indian monsoon mechanisms, so that a better informed attribution of change can be determined.

Generally, it is known that tropical depressions that form in the Bay of Bengal and then propagate westward or northward play a key role in extreme rainfall. These are associated with a mix of barotropic and baroclinic instabilities, and their interaction with the mean monsoonal flow(Gadgil [2003]). We suspect that the details of these interactions may be associated with the indicated changes in the spatial structure of the trends, and naturally for the trends themselves. A recent study by Guhathakurta and Rajeevan [2008] notes a significant decreasing trend in the frequency of depressions and storms over the Bay of Bengal, lending support to our conjecture. Model based studies and detailed analysis of individual extreme events are necessary to develop this intuition and are being pursued.

The design of our study was somewhat different from prior work. First, we considered changes in the climatology of extremes for each spatial location, rather than the frequency and intensity of exceedance of a fixed threshold, which is more meaningful for an analysis of the larger spatial scale considered. Second, instead of considering linear trends, we consider

the more general case of monotonic trends (this would include for example a step change in the process at some time, or an exponential or logarithmic trend), and assessed the evidence for such a trend using robust, nonparametric methods. The field significance test was applied both nationally and regionally and essentially confirms that the number of cases where the null hypothesis of no trend was rejected was statistically different from that obtained purely by chance (at the relevant level of significance).

Thus, our study lends credibility to previous assessments which report increasing trends in frequency and intensity of extreme rainfall over India, while identifying areas where there is a systematic departure from previous assessments. The issue of how these trends may be reinforced or reversed over the 21^{st} century is not addressed here, and is the subject of ongoing investigation. However, we note that the broad direction of trends identified here is consistent with the expectation from the model-based analysis for the 21^{st} century scenarios for climate change, as reported by Kumar *et al.* [2006] and May [2004].

Bibliography

Alexander,L.V., et al. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research*, 111:doi:10.1029/2005JD006290, 2006.

B. Bhaskaran and J.F.B. Mitchell. Simulated changes in southeast asian monsoon precipitation resulting from anthropogenic emissions. *International Journal of Climatology*, 15:1455–62, 1998.

A. C. Davison and D. V. Hinkley. *Bootstrap Methods and Their Application*. Cambridge Series in Statistical and Probabilistic Mathematics , No 1, 1 edition, 1997.

Kimberley L. Elmore, Michale E. Baldwin, and David M. Schultz. Field significance revisited: Spatial bias errors in forecasts as applied to the eta model. *Monthly Weather Review*, 134:519–531, 2006.

Sulochana Gadgil. The indian monsoon and its variability. *Annual Review of Earth and Planetary Science*, 31:429–67, 2003.

Xavier Gine, Robert Townsend, and James Vickrey. Statistical analysis of rainfall insurance payouts in southern india. *American Journal of Agricultural Economics*, 89:1248–1254, 2007.

B.N. Goswami, V. Venugopal, D. Sengupta, M.S. Madhusoodanan, and Prince K. Xavier. Increasing trends of extreme rain events over india in a warming environment. *Science*, 314:1442–1444, December 2006.

P. Guhathakurta and M. Rajeevan. Trends in the rainfall pattern over india. *International Journal of Climatology*, 28:1453–1469, 2008.

P. Hall and S.R. Wilson. Two guidelines for bootstrap hypothesis testing. *Biometrics*, 47:757–762, 1991.

Dennis R. Helsel and Robert M. Hirsch. *Statistical Methods in Water Resources*. Studies in environmental science, 49. Elsevier Science, Amsterdam, 1992.

U.R. Joshi and M. Rajeevan. Trends in precipitation extremes over india. Technical Report 3, National Climate Centre, 2006.

Klein Tank, A.M.G., et al. Changes in daily temperature and precipitation extremes in central and south asia. *Journal of Geophysical Research*, 111:doi:10.1029/2005JD006316, 2006.

K. Rupa Kumar, A. K. Sahai, K. Krishna Kumar, S. K. Patwardhan, P. K. Mishra, J. V. Revadekar, K. Kamala, and G. B. Pant. High-resolution climate change scenarios for india for the 21st century. *Current Science*, 90:332–345, 2006.

Murari Lal, Gerald A. Meehl, and Julie M. Arblaster. Simulation of indian summer monsoon rainfall and its intraseasonal variability. *Regional Environmental Change*, 1:163–179, 2000.

Robert E. Livezey and W.Y. Chen. Statistical field significance and its determination by monte carlo techniques. *Monthly Weather Review*, 111:46–59, 1983.

W. May. Simulation of the variability and extremes of daily rainfall during the indian summer monsoon and future times in a global time-slice experiment. *Climate Dynamics*, 22:183–204, 2004.

M. Rajeevan, Jyoti Bhate, J.D. Kale, and B. Lal. High resolution daily gridded rainfall data for the indian region: Analysis of break and active monsoon spells. *Current Science*, 91:296–306, 2006.

M. Rajeevan, Jyoti Bhate, and A.K. Jaswal. Analysis of variability and trends of extreme rainfall events over india using 104 years of gridded daily rainfall data. *Geophysical Research Letters*, 35:doi:10.1029/2008GL035143, 2008.

K.E. Trenberth. Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change*, 42:327–339, 1999.

D.S. Wilks. Resampling hypothesis tests for autocorrelated fields. *Journal of Climate*, 10:65–82, 1997.

Sheng Yue, Paul Pilon, and George Cavadias. Power of the mann-kendall and spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, 259:254–271, 2002.

chapter 4

Increasing frequency of extreme rainfall over Central India[1]

[1] Joint work with Upmanu Lall

4.1 Introduction

Increase in extreme precipitation events is a key concern of climate change studies. In the case of India, Rajeevan *et al.* [2008], Krishnamurthy *et al.* [2009], Goswami *et al.* [2006], Ajayamohan and A. Rao [2008] have attempted to analyze changes in the frequency and magnitude of extreme precipitation over India. All, except Krishnamurthy *et al.* [2009], focused on the Central India region referred to as the "core monsoon region"[2], to which we restrict the present analysis. Rajeevan *et al.* [2008] and Goswami *et al.* [2006] analyzed the data on frequency of exceedance of a specified threshold (100 and 150 mm) over the Central Indian region, and established a positive trend in the frequency at the higher threshold (150mm/day, which they label "Very Heavy Rain" or VHR events). They identify a significant connection between this measure of extreme and Equatorial Indian Ocean (EIO) SST. Ajayamohan and A. Rao [2008] consider the relationship between the frequency of extreme events (using the same threshold as in Rajeevan *et al.* [2008]) over a slightly different region (the Ganga-Mahanadi basin) and the DMI (Dipole Mode Index, an index indicating the strength of the Indian Ocean Dipole (IOD) phenomenon). They report both increasing trends and modulation of extremes by the IOD.

Past examination of trends in extremes of rainfall over Central India did not adequately account for the simultaneous influence of natural modes of climate variability (e.g. those related to ENSO or the IOD). The quasi-periodic and multi-scale nature of these climatic factors may compromise the results obtained from these analysis. Further, past analyses were mostly based on a shorter time series of 53 years of data ($1951 - 2004$), due to which it is also difficult to assess the significance of these other factors. In addition, the use of

[2]Defined as the region between 16.5-26.5 N and 74.5-86.5 E. The region analyzed in Rajeevan *et al.* [2008] was slightly larger than this, while that in Ajayamohan and A. Rao [2008] is a subset of this region. However, we will continue to refer to the regions used therein as being the core monsoon region except where we explicitly address the impact of these differences.

fixed rainfall levels to define thresholds in previous research does not lead to a consistent probabilistic threshold across the given region.

The current study relates selected climate indices (IOD, ENSO, EIO SST) and frequency of extreme rainfall events for the core monsoon (Central Indian) region, over a period of 97 years ($1901-1997$) and explores the potential connections between their respective cyclical and monotonic trends. Here, we consider a fixed probabilistic threshold across each grid box within the region, ensuring that we have a consistent probabilistic threshold across the domain, to assess exceedances. The analysis here is therefore more directly focussed on an assessment of whether there have been *changes in the probability of exceedances* over the given domain. We define the threshold in terms of the 99^{th} percentile of daily rainfall, and use the median of the 99^{th} percentile of daily rainfall, estimated for each year, as a *robust* estimate of the threshold.

This framework allows us to address the following important questions: What are the secular and quasi-periodic trends in the frequency of extreme rainfall events, and do these bear a relationship to teleconnection indices which have been conjectured to be important for this region ? In other words, this framework allows for a unified approach to testing for secular trends, in the presence of covariates, as well as for covariate significance.

4.2 Data

For rainfall, we use gridded ($1^0 \times 1^0$ resolution) daily rainfall data, from 1901-2004, over India, developed by the IMD and described in Rajeevan *et al.* [2008]. We use only the 133 grids in the region of interest (defined above) for our analysis. Gridded monthly data ($2^0 \times 2^0$) for the Equatorial Indian ocean region SST (EIO SST) was obtained from http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.ersst.html (accessed on June 8, 2010), for $1901-2004$; we use two definitions of the EIO, the first one as $5S-5N$ lat and $50E-100E$ longitude, while the second one (following Ajayamohan and A. Rao [2008])

is between $10S - 5N$ latitude and $60E - 90E$ longitude. For both definitions, we extract monthly data, compute the monthly anomalies for each of *JJAS* w.r.t the long-term (1901-2004) mean for each grid and subsequently compute the mean anomaly over the season across the grids.

We use the *Dipole Mode Index (DMI)* as a measure of the Indian ocean dipole (IOD), data for which are obtained from http://www.jamstec.go.jp/frsgc/research/d1/iod/(accessed on May 5,2010). We use the raw (without detrend and without time filter) monthly data, extract the data for the JJAS season and obtain the mean JJAS values. Data for the NINO34 and NINO12 indices were obtained from the IRI's data library (http://iridl.ldeo.columbia.edu/SOURCES/.I from which seasonal means for the JJAS season were computed for the years $1901 - 2000$.

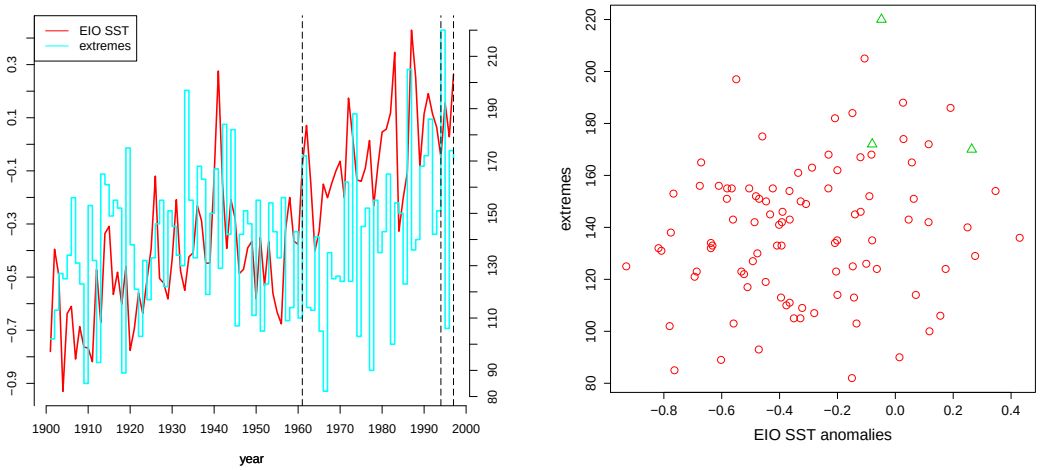
4.3 Analysis of extreme rainfall events

4.3.1 Definition of Threshold

We believe that a fixed choice of a threshold over heterogeneous regions is not appropriate. Instead, following Krishnamurthy *et al.* [2009], we define the threshold objectively, at each grid, as a fixed percentile of the long-term rainfall distribution at that grid cell. For each grid cell, the number of non-zero precipitation events each year was first identified and the 99^{th} percentile of this series was extracted (for each year). Subsequently, the *median* of these annual 99^{th} percentile across all years was selected to be the threshold for that particular grid cell, and rainfall events exceeding the threshold are taken to be extreme events (*frequency of extremes*). The spatially varying climatology of extreme rainfall across the core monsoon region is thus addressed. Finally, we obtain the sum of such events over all the grids for the core monsoon region each year, and thus, obtain a time series of extreme events.

From Figure 4.1(a), it is evident that there is significant heterogeneity in the extremes

of rainfall and choosing a fixed amount as the threshold, for instance 100 mm, will most likely yield a very large number of events, in certain grids, which are not in the tail of the probability distribution of rainfall, leading to a few grids dominating the count of extreme events in the region.



(a) Plot of EIO SST and extreme events

(b) scatter plot of EIO versus extreme events

Figure 4.1: (a) Line plots of EIO SST anomalies and extreme events. Dashed vertical lines correspond to IOD events at 1961, 1994 and 1997 (b) Scatter plots between Eq. Ind. Ocean SST and frequency-of-exceedance at lower and higher thresholds, respectively. Red triangles in both correspond to the aforementioned IOD events.

Figure 4.1(a) also indicates that there has been an increase in both SST anomalies and the frequency of extremes over the latter half of the 20^{th}century. However, it is not clear that there is any particularly significant association between them, as is clear from Figure 4.1(b). Further, if the conjecture that strong IOD anomalies modulate more extreme events (as in Ajayamohan and A. Rao [2008]) is true, then years with very strong IOD signals (such as 1961, 1994 and 1997) must also be associated with more than usual extremes. From Figure 4.1(b), no such consistent pattern can be discerned, apart from the extremely strong IOD event in 1994.

Many of the prior analyses seek to develop such relationships after detrending (or otherwise filtering) the series. Recognizing the structured low-frequency variability in the climate indices and in extreme rainfall is also an issue, we explore the trends and teleconnections in the frequency domain.

4.3.2 Changes in the relationship between frequency of extremes and climate indices: Spectral analyses

4.3.2.1 Multi-taper spectral analysis

We investigate fluctuations in the frequency of exceedance [3] using the Multi-taper method of spectral analysis. This method provides a more robust characterization of narrow-band variability than do traditional methods of spectral analysis. We refer the interested reader to Mann and Lees [1996], especially their Appendix A, for more details, and to Lall and Mann [1995], for a detailed application.

A summary of the individual spectra as well as coherences between the individual series and relevant climate indices are presented in Table 4.1. A few of the important points which arise from an inspection of the table are presented below.

Figure 4.2(a) is the MTM spectrum with "robust" confidence intervals at the lower and higher thresholds, respectively[4]. Two "significant" peaks are evident from the figure, that near the 0 frequency and at approximately the $ENSO$ frequency. In particular, the only

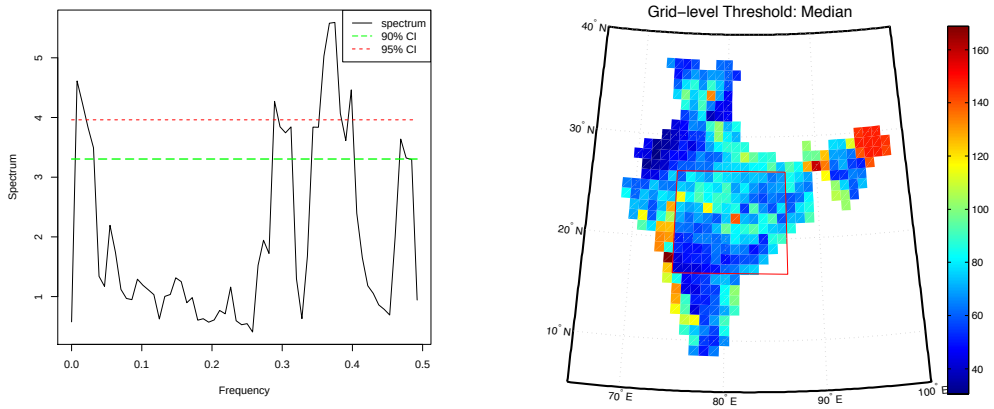
[3] There are two distinct uses of the term "frequency" here; the first referring to the "frequency of exceedances of extreme rainfall" and the second to the frequencies in the frequency-domain analysis; in order to avoid confusion in usage, henceforth, we refer to the count of extreme events as simply "exceedances", instead of "frequency of exceedances" (as used up to now). Therefore, any further references to frequency, unless otherwise indicated, shall refer to frequency domain analysis.

[4] Denote by K the number of tapers used, by W the half-bandwidth, by $\Delta t = 1$ (in annual units) the sampling frequency, by $f_R = \frac{1}{n * \Delta t} = \frac{1}{97 * 1} = 0.0103$ the Rayleigh frequency. Writing, as is usual, $W = pf_R$ setting $p = 2$, the number of tapers now is $K = 2p - 1 = 3$ and the resulting DPSS is called a 2π-taper. The spectral estimate thus averages in the frequency band $f \pm pf_R = f \pm 0.0206$, where we will term pf_R as being the minimum resolvable frequency. Thus, we will only speak of peaks which are "well separated" from each other, in the sense of being at least $2pf_R$ distance apart. See for instance [Lall and Mann, 1995, pp 2510]

Individual Spectra		Coherence between exceedance counts and	
series	Significant periods (years)	series	Significant periods (years)
Extreme events	*trend*, $3.25-4, 2.25-3$	NINO12	decadal, $3.5-4$
NINO12	$5-5.5, 3.5-4$	NINO34	$7-10, 3.5, 2.5-3$
NINO34	$6.5-5.25, 3.5$	EIO SST	decadal, $3, 3.5, 2.5$
EIO SST	trend, 5	DMI	$3.75, 3$
DMI	trend, $4.75-5.5$		

Table 4.1: Results of Multi-taper Spectral Analysis based on Mann and Lees [1996]. Significance (at 95% level) of frequency bands (or individual frequencies) are based on an inspection of the (*jacknife*) confidence intervals. We only consider frequencies which are "well separated" from each other (see footnote 4). Features near the 0 frequency are referred to as "trends", but note that it is not possible to distinguish multi-decadal cycles and processes with long memory from those with a time trend.

feature longer-than-decadal of interest is the time trend. The results here are consistent with the analysis in Krishnamurthy et al (2010), who undertake a more comprehensive multivariate frequency-domain analysis of trends in frequency and intensity over India using the same data set, and find four different time scales of interest, the time trend, *ENSO*, *QBO* and the *decadal* signal, of which the *decadal* signal is the weakest.



(a) extreme events (b) Spatial distribution

Figure 4.2: (a) MTM spectrum of the frequency-of-exceedance data. Minimum resolvable frequency is $pf_R = 0.0357$, see footnote 4 (b) Spatial distribution of the thresholds chosen to define extreme events

The spectrum of the EIO SST series, in Figures 4.4(b) and 4.4(c), indicates clearly that power exists in only two frequency bands, the secular trend and the QBO[5]; further, the rapid decay of the spectrum suggests a predominance of low frequency (greater than 5 years) components. The NINO12 series, Figure 4.4(a) appears to possess significant power at an approximately $5 - 5\frac{1}{2}$ year band and a $3\frac{1}{2}$ year band; the DMI series, Figure 4.4(d) has significant power at the secular trend and a QBO-like $4\frac{3}{4} - 5\frac{1}{2}$ year band. In particular, we note that none of the climate indices exhibit any significant power at the decadal scale or lower (apart from the secular trend), and this is very similar to the findings in the case of the frequency of extremes at both the thresholds.

In order to assess the common evolution between the individual series and Eq. Indian ocean SST, the DMI, we plot the MTM squared-coherence between these series and exceedance counts. From the Figures in 4.3, the following are evident:

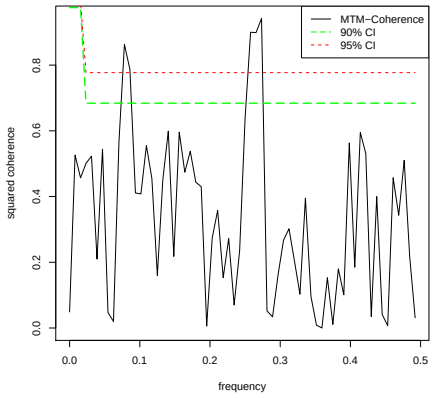
[5]An advantage of the MTM spectrum over the global wavelet spectrum is that it is better suited at identifying irregular phenomena such as the QBO. See for instance Torrence and Webster [1999], who indicated that the Morelet wavelet, in particular, was not very suited to identifying the QBO.

1. The decadal signal appears to be significant with all climate indices; however, noting the low power at this frequency range in all climate series as well as the frequency series, we consider it an artifact of the procedure and do not further analyze it.

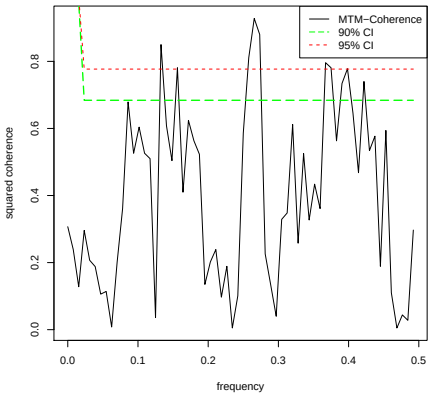
2. With NINO12 and NINO34, significant relationships appear to exist at a time period centered around $3\frac{1}{2}$ years.

3. With EIO, the only relevant significant relationship is at the secular trend and at an ENSO-like band of $3\frac{1}{3} - 4$ years.

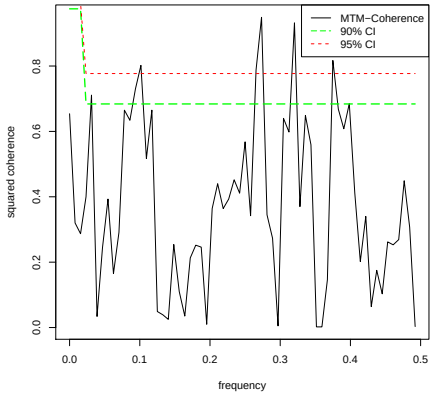
4. Significant frequencies with the DMI appear to be at an ENSO-like 3-4 year band (where both spectra have significant power) .



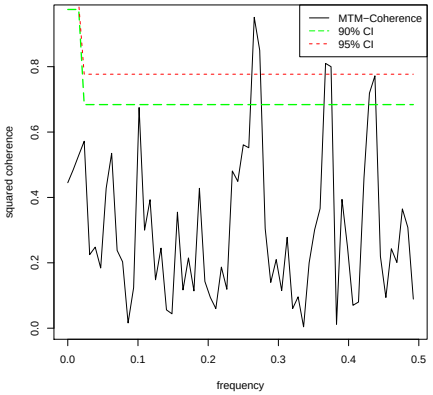
(a) NINO12



(b) NINO34



(c) EIO



(d) DMI

Figure 4.3: (Squared) Spectral Coherence between frequency of exceedance and various climate and SST indices (indicated in the respective figures). Other conventions same as for figure (4.2)

4.3.3 Trends and relationship to climate indices

In much of the literature related to changes in frequency of extremes in precipitation, typically two exercises are pursued disjunctively:

1. testing for existence of (monotonic, generally linear) trends in extremes of rainfall

2. exploring the impact of covariates on extreme rainfall.

We illustrate that, in the presence of covariates which potentially impact the series under consideration, tests for trend must account for the impact of such covariates[6].

	Quantile Regression				Poisson Regression	
	0.25 Q	0.5 Q	0.75 Q	0.9 Q		
NINO12	-0.0162	0.0188	-0.0316	0.0538***	-0.02	
	(0.887)	(0.681)	(0.670)	(1.83e-05)	(0.24)	
Dipole Mode Index	0.0176	0.0334	-0.0030	-0.0467*	-0.007	
	(0.919)	(0.637)	(0.984)	(0.0759)	(0.82)	
Eq Indian Ocean SST	-0.0062	-0.0780	0.2192	0.2136***	0.111*	
	(0.965)	(0.217)	(0.600)	(0)	(0.02)	
time	0.0006	0.0022***	-0.0000	-0.0004	0.001	0.001***
	(0.766)	(0.000731)	(0.999)	(0.412)	(0.14)	(0.00)

p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.2: Results for Poisson and Quantile Regressions.

Consider first the results of the canonical model for count data, the Poisson regression model (details of which maybe obtained from Winkelmann [2008](Chapter 3)), in the final two columns of Table (4.2). Two important conclusions emerge from a perusal of the table: in the presence of other covariates, time trends do not appear to be significant and second, in a model with only time, the coefficient on time is significant, indicating that a time trend is, indeed, present. It is therefore apparent that the presence of a trend is not robust to inclusion of other covariates; for instance, the EIO SST series itself appears to have a time trend (see for instance the correlation with time in Table (3)), and it is likely, as conjectured in Rajeevan *et al.* [2008], that there is a relationship between the two. It is also likely that the trend in the EIO SST is a result of global warming. Thus, it is likely that what is labeled a "time trend" is simply a relationship with EIO SST (which is possibly trending).

[6]Given that we focus on exploring the full conditional distribution of frequency of extremes, we do not use the many proposals for trends tests for the mean of a count variable, such as the parametric ones in Frei and Schär [2001]; Keim and Cruise [1998] as well as the nonparametric one in Brillinger [1994].

A mean regression, however, is unsatisfactory in the current context. For instance, we are interested in understanding the variation in the areal extent impacted in terms of threshold crossings; it is not clear, in this case, that the mean effect captures sufficient spatial heterogeneity. In other words, if covariate impacts are evident only in larger scale events, the conditional mean framework–with a uniform threshold–is likely to provide only very incomplete information. Further, even in the case of only a single grid box, we anticipate that covariates will exert differential impacts on strong events.

Quantile regression is a natural framework to accommodate heterogenous impact of covariates at different points along the conditional distribution; for instance, a popular interpretation of quantile regression is that the covariates affect not only the location but also the scale (variance) of the conditional distribution of the response. In the current context, this framework would allow us to disentangle two issues: covariate significance (including time trends) and heterogeneity of impacts. Given the discrete ("count") nature of our response variable (frequency of extreme rainfall events) we use a newly developed framework for quantile regression for count data, following Machado and Silva [2005]. Details on various types of quantile regressions maybe found in Koenker [2005] while instances of the application of quantile regression for climate data are Elsner *et al.* [2008]; Koenker and Schorfheide [1994]. Ours appears to be the first instance of usage of quantile regression for count data in the climate literature, and therefore, we provide a slightly detailed introduction to the method in Appendix C.

To summarize, quantile regression allows us to evaluate if climate teleconnections to exceedances over the Central Indian region vary at the large and small scale (translated into higher and lower conditional quantiles). We report results, in Table (4.2), on the first three quartiles and in addition, at the 90^{th} percentile, a common choice.

Two important features are evident from the table; time trend is significant and positive at *only* the median [7], and climate tele-connections are evident only at the 90^{th}percentile. We now assess the impact of different climate indices separately, in turn. We find that there is a significant modulation of count of exceedances by the DMI but only at the highest quantile considered, 90^{th} percentile. Finally, EIO SST and ENSO are both associated with an increase in the count of exceedances but again, at only the 90^{th} percentile. We may summarize the results of quantile regressions as follows[8]: time trends are significant at only one part of the distribution (the median); the DMI modulates count of exceedances only at the 90^{th}percentile. EIO SST and ENSO both appear to be associated with very heavy rainfall events. In general, impact of the covariates appears to be different at different quantiles, and especially at the 90^{th} percentile (at which there is insufficient data to permit a more serious examination).

Two important aspects of the results are worth discussing: the negative impact of the DMI and the positive impact of ENSO. Unlike Ajayamohan and A. Rao [2008], our findings indicate that, over sufficiently large regions of Central India, the impact of IOD events (which appear as positive anomalies in the DMI and as negative anomalies in the Eastern

[7] Note that the model is exponential in the parameters, and therefore, time trends–as well as the effects of the other covariates–here are not linear; however, when the coefficients are very close to zero, as the is the case for time trends, they may effectively be treated as being linear. This is not however the case with other coefficients since they are not very close to zero, and in these cases, conventionally, in addition to coefficients, reports are provided of "marginal effects" i.e. $\partial Y/\partial X_k$, the derivative of Y w.r.t to the K^{th}covariate. In this particular model, we have that: $\frac{\partial Y}{\partial X_K} = \beta_K exp(\beta X)$; typically, the effects are reported for each covariate with all covariates held at their median values. We do not report these since in the current case, it is not clear that such marginal effects convey much more information than do the coefficients. However, these marginal effects are available from the authors upon request.

[8] The significance of the time trend in the Poisson regression depended on the specific definition of the EIO SST index used; further, the significance of the time trend does not survive robust covariance matrix or the usage of more realistic models for conditional mean of count data, such as the negative binomial model. We thus wish to emphasize that the results are not an artifact of covariate selection. In addition, even the popular, simplistic alternative to Poisson regression, wherein $log(y)$ is regressed on covariates, led to very similar sensitivity of significance of trend on definition of the EIO SST used, and in general, yielded very similar results. The quantile regression results, however, proved robust to changes in the definition of the EIO SST as well. In addition, the effects of DMI also proved robust to changes in definition of the EIO SST and these results are not reported for brevity. Detailed results are available from the authors upon request.

DMI (EDMI) anomalies used in Ajayamohan and A. Rao [2008]) is negative i.e. positive IOD events lead to *reductions* in the spatial extent of exceedances, once the impacts exerted by other important tele-connections such as ENSO are accounted for. Further, ENSO and EIO SST both exert positive impacts indicating that tele-connections to both are strong. Rajeevan *et al.* [2008] indicate a strong impact of both indices on extreme rainfall events over Central India, individually; our results clarify that they both jointly impact extreme events over the same region, and further, that the relationship is significant only over a sufficiently large spatial scale.

We emphasize that we do not draw serious inferences about changes in signs and significance at extreme quantiles ; rather, we offer the evidence above as an alternative and more comprehensive view of the relationship between climate covariates and frequency of extreme rainfall.

4.4 Summary and conclusion

This study extends earlier work on extremes of monsoon rainfall over Central India in several directions. First, using a more appropriate measure of extreme rainfall, a comprehensive analysis of the dominant modes of variability and their variation over time was carried out using multi-taper spectral analysis. It was shown that most of the power in the count of exceedances (at both thresholds) was concentrated in the *ENSO* and the *QBO* bands. Further, coherence in the evolution of frequency of extremes and climate indices indicated to exert influence on it (EIO SST, DMI and ENSO) was investigated, and it was shown that:

- coherence with ENSO exists in the 3-4 year band
- coherence with the EIO SST exists only at the secular trend and the 3 1/3 -4 year ENSO-like band

- coherence with the DMI is at the 3-4 year band, along with some higher frequency components.

Second, prior analyses of the impact of climate indices and existence of time trends was extended to different quantiles of the distribution of the counts of exceedances. Results indicate that time trends were possibly not significant except at the median of the distribution. Further, the impact of EIO SST is observed only at the upper tail of the distribution of exceedances. It was also shown that the IOD exerts a modulating influence on count of exceedances, only at the 90^{th} percentile. Finally, impact of ENSO on frequency of extremes were also noted only at the 90^{th} percentile.

The non-uniformity of the impact of EIO SST on different quantiles of the frequency of exceedance as well as its lack of coherence at most frequencies suggest a relationship between them which is far from a simple coupling. The implications for flood frequency of (model) projected increases in EIO SST are therefore somewhat unclear. More puzzling are the (physical) causes of modulation of frequency of extremes by the DMI, as well as the (empirical) weakness of the relationship between them, despite coherence over a range of frequencies. Ajayamohan and A. Rao [2008] offer a reasonable physical explanation for the modulation by DMI, but one that is not yet empirically supported.

We do not offer any physical mechanism for the influence of either EIO SST or DMI on frequency of extremes over Central India but note that in the absence of either a simple coupling mechanism (as with the EIO SST) or an empirically verifiable physical mechanism (as with the DMI), conclusions of enhanced SST or frequency of IOD events leading to increased extremes over Central India are, at best, conjectures.

Bibliography

R.S. Ajayamohan and S. A. Rao. Indian Ocean Dipole modulates the number of extreme rainfall events over India in a warming environment. *Journal of the Meteorological Society of Japan*, 86(1):245–252, 2008.

D.R. Brillinger. Trend analysis: time series and point process problems. *Environmetrics*, 5(1):1–19, 1994.

J.B. Elsner, J.P. Kossin, and T.H. Jagger. The increasing intensity of the strongest tropical cyclones. *Nature*, 455(7209):92–95, 2008.

C. Frei and C. Schär. Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. *Journal of Climate*, 14:1568–1584, 2001.

BN Goswami, V. Venugopal, D. Sengupta, MS Madhusoodanan, and P.K. Xavier. Increasing trend of extreme rain events over India in a warming environment. *Science*, 314(5804):1442, 2006.

A. Grinsted, J.C. Moore, and S. Jevrejeva. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11(5/6):561–566, 2004.

B.D. Keim and J.F. Cruise. A technique to measure trends in the frequency of discrete random events. *Journal of Climate*, 11:848–855, 1998.

R. Koenker and F. Schorfheide. Quantile spline models for global temperature change. *Climatic Change*, 28(4):395–404, 1994.

R. Koenker. Quantile regression. Econometric society monograph series, 2005.

C.K.B. Krishnamurthy, U. Lall, and H.H. Kwon. Changing Frequency and Intensity of Rainfall Extremes over India from 1951 to 2003. *Journal of Climate*, 22:18, 2009.

U. Lall and M. Mann. The Great Salt Lake: A barometer of low-frequency climatic variability. *Water Resources Research*, 31(10):2503–2515, 1995.

J.A.F. Machado and J.M.C.S. Silva. Quantiles for counts. *Journal of the American Statistical Association*, 100(472):1226–1237, 2005.

M.E. Mann and J.M. Lees. Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, 33(3):409–445, 1996.

A. Miranda. Planned fertility and family background: a quantile regression for counts analysis. *Journal of Population Economics*, 21(1):67–81, 2008.

M. Rajeevan, J. Bhate, and AK Jaswal. Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. *Geophys. Res. Lett*, 35, 2008.

C. Torrence and P.J. Webster. Interdecadal changes in the ENSO-monsoon system. *Journal of Climate*, 12(8):2679–2690, 1999.

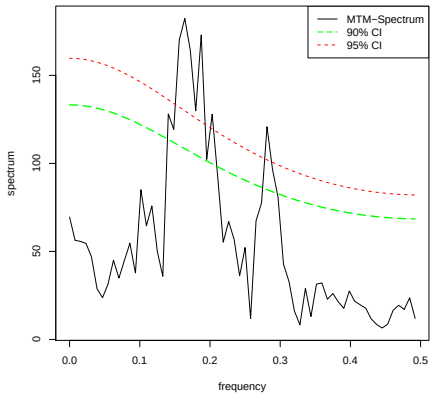
R. Winkelmann. *Econometric analysis of count data*. Springer Verlag, 2008.

Appendix 4.A Details of software used

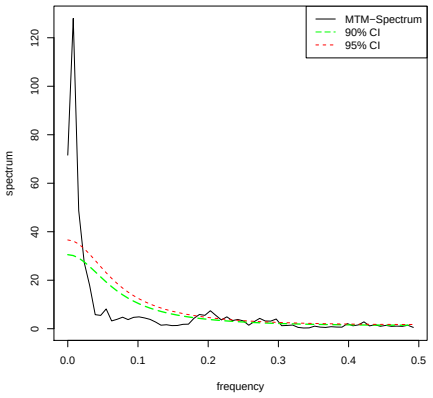
For MTM spectral analysis, we used Mann and Lees [1996] fortran package, modified for our version of Linux, while for wavelet analysis, we used the matlab package provided by Grinsted *et al.* [2004], downloadable from http://www.pol.ac.uk/home/research/waveletcoherence/ (downloaded on June 10, 2010). For Quantile regression, we used the freely downloadable add-on QCOUNT to STATA program, provided by Miranda, A Miranda [2008]. All programs (in fortran, Matlab and Stata) are available upon request from the author, along with the full dataset used, to facilitate reproduction of all the results. Similarly, all the figures and tables referenced but not presented in the text are also available from the authors upon request.

Appendix 4.B Supplemental Figures and Tables

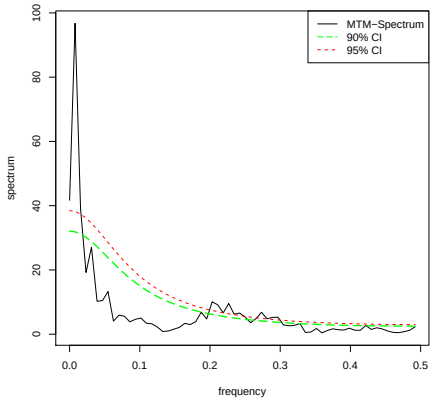
Multi-taper spectra of various climate indices.



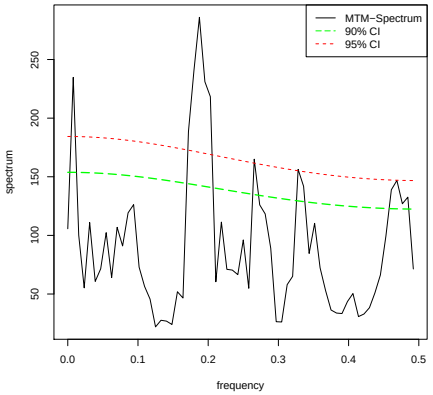
(a) NINO12



(b) EIO SST



(c) EIO2



(d) DMI

Figure 4.4: MTM spectra of various climate indices. Confidence intervals are as for figure (4.2).

Table of correlation coefficients between various covariates

Appendix 4.C Quantile Regression for Count Data

The discussion here follows the simplified one in Section 6.6 of Winkelmann [2008], to which the reader is directed for a slightly more thorough (and accessible) discussion. Denote Y as

	NINO12	NINO34	EIO	DMI
DMI	0.260552	0.148625	0.281787	
	[0]	[0.03]	[0]	
EIO	0.25196	0.203608		
	[0]	[0.003]		
Freq	0.012	-0.006		0.0085
	[0.86]	[0.993]		[0.9]
Time	0.07	0.026	0.505	0.185
	[0.274]	[0.7]	[0]	[0.008]

Table 4.3: Kendall's rank correlation between different predictors (p-values in brackets)

the dependent variable, X as the matrix of $K+1$ covariates, (including a constant), and β the coefficient on a regression of Y on X. Inference on the coefficient β in a quantile regression setting depends on the *continuity* of the conditional quantile function at the τ^{th} quantile, commonly denoted $Q_\tau(Y/X)$. When the dependent variable Y is a *count* variable i.e. when it can take only a finite number of values from (a finite subset of) $\mathbb{N}$, this assumption is untenable. Machado and Silva [2005] surmount this obstacle by constructing a *continuous* random variable, Z, whose quantiles have a bijective relationship to that of the variable of interest Y. In particular, letting U be a vector of uniform random variables, they define $Z := Y + U$ and induce smoothness into the problem, leading to a conditional quantile which is continuous. They further note that there exists a simple monotonic relationship between the quantiles of Z and that of Y, which implies that the significance of regressors in $Q_\tau(Y/X)$ maybe tested by means of their significance in $Q_\tau(Z/X)$.

In order to ensure non-negative quantiles, one can define the quantile of the transformed version of the random variable, where the transformation will be locally (at the quantile of interest) positive. Thus, the quantile they work with is $Q_\tau(T(Z;\tau)/X) = X\beta$, where $T(:,\tau)$ is a monotonic transformation of Z, whose inverse exists. The particular inverse transform used in Machado and Silva [2005] is defined by:

$$Q_\tau(Z/X) = \tau + exp(X\beta)$$

Thus, $Q_\tau(Z/X)$ is bounded below by τ and is non-negative[9]. Under relatively mild conditions, which includes the important assumption that ***at least one*** of the covariates is continuous, they indicate the following framework for quantile regression with count data:

1. Obtain $Z = Y + U$

2. Estimate the quantile regression of

$$T(Z;\tau) = \begin{cases} \log(Z-\tau) & \text{for Z}\geq \tau \\ \log(\zeta)\text{for} & Z \leq \tau \end{cases}$$

   on X (as with the continuous dependent variable case), with ζ typically being a very small number (say 10^{-4}).

3. Carry out inference on the resultant quantiles, using the monotonic relationship between $Q_\tau(Z)$ and $Q_\tau(Y)$.

4. In order to not base inference on a specific realization of the uniform random variable, the coefficient estimates and covariance matrices are averaged over m draws (we use $m = 50,000$ which is much larger than the $10,000$ recommended by Machado and Silva [2005])

[9] In more detail, the count variable is by definition non-negative, in which case simply writing $Q_\tau(Z/X) = X\beta$ may well lead to estimates of β for which the predicted value of Y is *negative*. In order to surmount this issue, in the case of mean regression (say the usual poisson regression), one models the conditional mean, conventionally called λ, as being exponential in the covariates i.e. the specification usually followed in poisson regression is: $\log \lambda_i = X_i\beta$, where $\lambda_i$is a the mean of the count process Y_i. In the case of quantile regression, therefore, a similar strategy can be used, especially since $Q_\tau(Z/X)$ is bounded below by τ, which motivates the exponential form with the additive term τ.

chapter 5

Spatio-temporal patterns and Teleconnections of Extreme Rainfall events over India[1]

[1] Joint work with Upmanu Lall

5.1 Introduction

Previous analyses of changes in patterns of extreme rainfall for India (Goswami *et al.* [2006]; Krishnamurthy *et al.* [2009]; Rajeevan *et al.* [2008]) or across the world (Groisman *et al.* [2005]; New *et al.* [2006]; Peterson *et al.* [2008]; Schmidli and Frei [2005]; Zhai *et al.* [2005]) have focused mostly on deterministic trends. However, in an analysis of extremes, the questions of interest often include:

1. Are there any deterministic trends?
   (a) If so, are such trends spatially coherent?
   (b) Can these trends be related to any global drivers of change, such as ENSO etc?
2. What are the dominant spatial and temporal modes of behavior of extremes?
3. Do extremes exhibit cyclical behavior?
4. Are there any precursors (in terms of SST field etc) to extreme events?

While past analyses of trends have been limited to answering question(1) above, this chapter reports on an ongoing study attempting to address many of the questions posed above.

Using a daily gridded ($1^\circ \times 1^\circ$) precipitation data set for $1901 - 2004$ from the IMD (Rajeevan *et al.* [2008]), and following on from Chapter 1, this study seeks to address the following important issues:

- Are there monotonic (deterministic) trends in extreme rainfall when a century of data are considered?
- If so, which regions experience such a trend? How do they differ compared to the shorter data set?
- What can be learnt if a simultaneous analysis of secular and quasi-oscillatory phenomena in extremes are carried out?

We report on the progress made in addressing these issues and briefly outline tentative conclusions, and provide indications for future work.

5.2 Methodology

5.2.1 Trend Analysis

For analyzing the trends, we follow the methodology in Krishnamurthy *et al.* [2009] i.e. we first apply two thresholds (see reference for definition of threshold and its application) and obtain the frequency (number of exceedances per year) and intensity (normalized total rainfall when the threshold is exceeded). We then apply, at each grid, the MK test procedure to detect monotonic trends.

5.2.2 Accounting for multiple tests

Given that there are a total of 357 simultaneous tests of monotonic trends, with not all tests being independent, we are interested in two issues (given that climate data exhibit strong spatial similarity in behavior):

1. Is there any spatial signal in the grids which exhibit a trend (i.e. does the "field" as a whole exhibit significant trends)?

2. If so, which are the regions which exhibit such a trend?

This is the problem of "test multiplicity" in the Climate literature (also called "multiple testing" in the Statistics literature). Following Ventura *et al.* [2004], we use the procedure of "False Discovery Rate" (explained below) to address the second issue. Issue (1) above, as pointed out in Wilks [2006], is the same as (2) and so we do not deal with (1) separately.

5.2.2.1 False Discovery Rate for independent tests

Procedure

Let K denote the total number of *independent* tests, τ_i the test statistic at location i and consider the test of interest for each grid (we consider below a two-sided test and note that everything carries over to the one-sided test, with obvious changes in critical regions for the test):

$$H_0 : \theta = \mu \tag{5.1}$$

$$H_a : \theta \neq \mu \tag{5.2}$$

The test is carried out at each of the K locations and interest lies in determining which of the tests are "significant" (i.e. interest lies in determining local significance). Denote the ordered $p-values$ from the tests $p_{(i)}$ (with unordered $p-values$ being denoted p_i) let α be the nominal probability of false rejections of the null, henceforth referred to as the *False Positive Rate* (or the probability of a type I error). Denote by $n_{H_0}$the total number of correct null hypothesis, n_{reject} the number of null hypotheses rejected and by n_{FN} the total number of false rejections of the null (or false positives), all of which are unknown (random variables).

The False Discovery Rate (whose nominal value is commonly denoted q) is defined as the expectation of the proportion of false rejections out of total rejections *i.e.*

$$FDP = n_{FP}/n_{reject}; FDR = \mathbb{E}(FDP) \tag{5.3}$$

while the False Positive Rate is defined as

$$FPP = n_{FP}/n_{H_0}; FPR = \mathbb{E}(FPP) \tag{5.4}$$

In a traditional setting, the null hypothesis is rejected at location i if $p_i \leq \alpha$ with the same α used for all locations. In the FDR method of Benjamini and Hochberg [1995] (BH, henceforth), however, the null is rejected if $p_i \leq p_k$ where

$$k = max_{i=0,1,\ldots,N} \left\{ i : p_{(i)} \leq \frac{qi}{n} \right\} \tag{5.5}$$

The null is therefore rejected at all locations where the $p - value$ is below a particular threshold (see Fig 2 and the accompanying discussion in Ventura *et al.* [2004]).

Rationale

Both FDR and FPR control the number of false detection, n_{FN} but in different ways. In the traditional procedure, the number of false detections (i.e. false rejections of the null) are proportional to the number of true null hypotheses and therefore, when a large number of null hypotheses are true, there are a large number of false detections. Corrections to α to limit false detections, usually by reducing α either in an ad-hoc manner or through a sample-dependent manner, as in the case of the *bonferroni correction*, reduce the probability of detecting changes and increases n_{FN}, (number of false negatives).

The FDR (for independent data) of *BH* controls the FDP in a meaningful way irrespective of the number of true null hypotheses i.e. $FDR = \mathbb{E}(FDP) \leq q$, independent of n_{H_0}(choice of q involves the same trade-off involved in a choice of α). There are two advantages to using the FDR (over the FPR):

1. The FDR is a more tangible and informative criterion of error control

2. Since n_{reject} is known (in the sample), an approximate upper bound on the average number of false detections is obtained via $\mathbb{E}(n_{FP}) \leq qn_{reject}$. Therefore, for any

sample of tests, we may always report, for a given q and estimated n_{reject}, that on average, there are fewer than qn_{reject} number of false positive detections (unlike in the traditional tests, since such a number cannot be reported without a knowledge of n_{H_0}, which is in general not estimable; see however, the section on correlated tests).

Ventura et al. *[2004]* propose a modified version of the test for correlated data, in particular, for climate data; however, as already noted in Wilks [2006], this procedure suffers from the drawback that, in many instances, it estimates a (possibly unacceptably) high rate of null rejections[2]. We therefore use the original procedure recommended by *BH* since, as already noted (in Ventura *et al.* [2004], among others), this methodology is surprisingly valid under general dependence.

A drawback of controlling the FDR, which we use here, is that most of the statements pertain to the mean of the climatic data generating process, of which the given data set is a realization, rather than a statement regarding the data set under consideration[3]. Thus, for instance, claims such as "the proportion of false rejections of the null are less than q" cannot be made; what can actually be said is "asymptotically, the mean proportion of false rejections are less than q". This distinction is important to note, since even when the FDR is controlled at q, the FPR may be very much higher than q (see Korn *et al.* [2004] for a realistic non-climate instance of this issue)

The focus, in the present analysis, is on an identification of the location of grids exhibiting trends, rather than tests of significance of "field"[4]. Ignoring, therefore, the issue of field

[2]In our case, using this procedure, for frequency of exceedance exhibiting positive significant trends, at the 99^{th}%tile, we see that the number of grids deemed significant is more than that at the nominal FPR; while possibly valid, without further research on the properties of this modified procedure for arbitrary correlation (not just the positive one for which it was validated in Ventura *et al.* [2004]), it is difficult to put more credence in this particular methodology

[3]In most climate applications, what is of actual interest is the FDP, since there are never many realizations of a dataset. However, given the relative lack of research in control of FDP, we resort to controlling the FDR, which is an expected value.

[4]Nonetheless, we point out that, in the present setting, Wilks [2006] shows that, while the FDR was developed for obtaining evaluations of local significance, it may also be seen as a test of the *global* or *field* significance. In fact, and very simply, a global null hypothesis, at level α_{global} is rejected if *at least one*

	Positive trends		Negative trends	
	FPR	**FDR**	**FPR**	**FDR**
frequency (90^{th} **%tile**)	55	1	54	27
frequency (99^{th} **%tile**)	64	16	37	14
Intensity (90^{th} **%tile**)	66	13	35	13
Intensity (99^{th} **%tile**)	64	8	35	0

Table 5.1: Number of grids with positive and negative significant frequencies (as established by the MK trend test); in each case, the relevant one-sided test, at the 10% level, was applied; "FPR" refers to the conventional procedure while "FDR" refers to the *BH* procedure

significance (in fact, subsequent spatio-temporal analysis via the MTM-SVD methodology maybe viewed as a more robust method of obtaining "field significance"), we focus on determinations of local significance, and plot below, each of the plot indicating "significant" grids using the conventional (or "FPR") method, the locations at which significance has been determined using the *BH* method.

We report most of the results of the testing procedures above as plots, to aid comprehension. We plot

1. all significant (at the 10% level) grids
2. all positive significant (at the 10% level) grids
3. all negative significant (at the 10% level) grids

for all measures of extremes we consider.

5.2.3 Trends in Frequency and intensity

From Table 5.1, we already observe a pattern emerging (with the FDR), in that, at the lower threshold, negative trends (weakly) pre-dominate while at the higher thresholds, positive trends are more prominent; however, this is not consistent across FPR and FDR. This already lends credence to the hypothesis that trends are indeed increasing at very high thresholds. There is also another aspect regarding FDR procedure in Table 5.1 which is worthy of comment: no grids are significant, when considering negative trends only, in the case of intensity at the 99^{th}percentile. This indicates that, at the higher threshold, there is no decrease in the intensity of extreme rainfall, a fact of some importance, which would have been completely unnoticed when applying the conventional test of significance (which indicates as many as 35 negative trends) or conventional field significance tests (which may well have indicated field significance). In fact, even if such a field significance test had indicated a non-significant field, it would not still have been possible to have surmised that not a single grid would have been significant (all of these interpretations are with the caveat in the previous section).

Figures 5.1 to 5.4 indicate the results of applying the MK trend test; panel 1 of each figure indicates all grids deemed significant using the same significance threshold (10% was used uniformly) while panel 2 indicates all grids deemed significant by the *BH* procedure outlined above.

---

local null hypothesis is, at the nominal level $q = \alpha_{global}$ (by any of the *FDR* procedures outlined in Section 5.2.2.1 above). It is not however, clear that the FDR control is the correct approach to testing for "field significance" since control of the FDR does not imply control of the FDP, due to which we cannot make claims about the *actual* (as opposed to *expected*) fraction of local null hypotheses rejected (asymptotically), which is what "field significance" attempts to report.

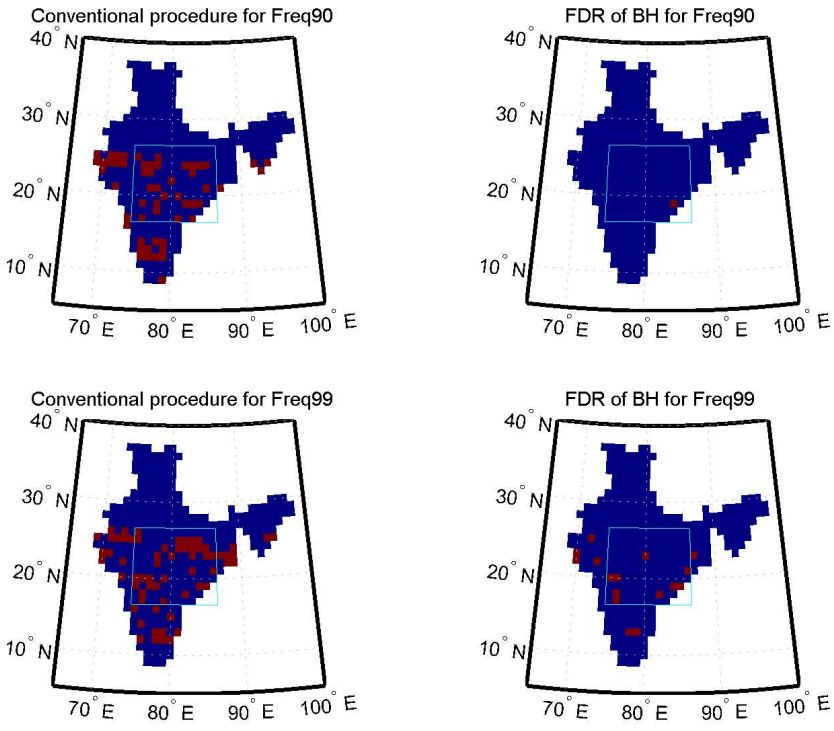


Figure 5.1: Spatial locations of the grids deemed to have positive significant trends in frequency of exceedance using the conventional and the BH procedures; the boxed area is the "core monsoon region"

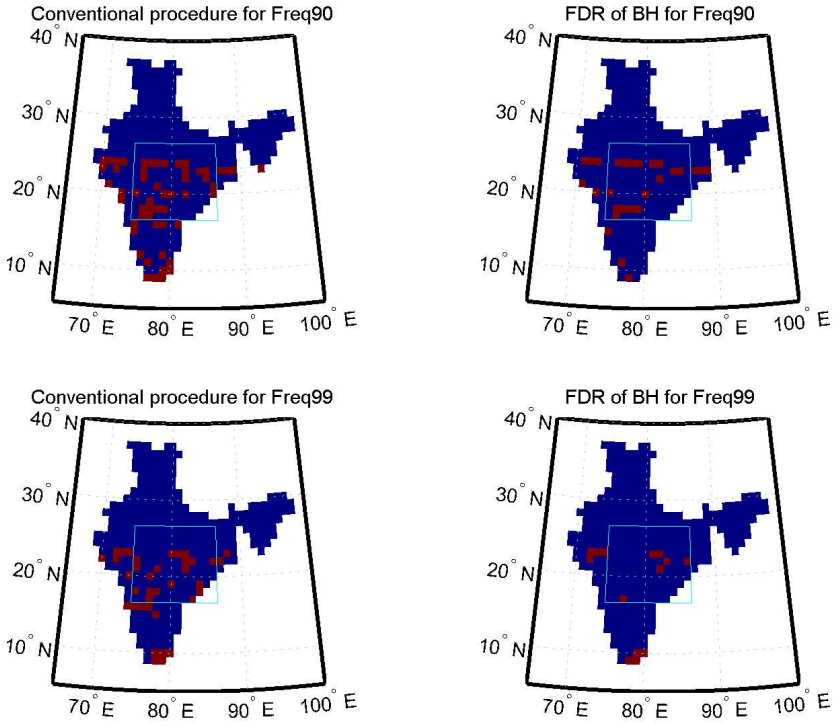


Figure 5.2: Same as Figure 5.1 but for grids with negative trends

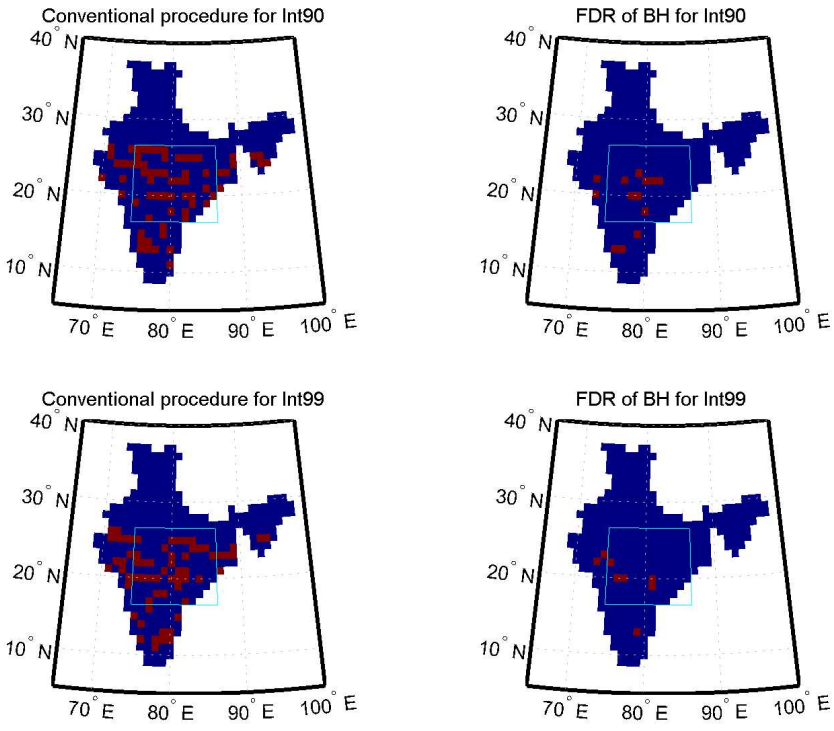


Figure 5.3: Spatial locations of the grids deemed to have positive significant trends in intensity of exceedance using the conventional and the BH procedures; the boxed area is the "core monsoon region"

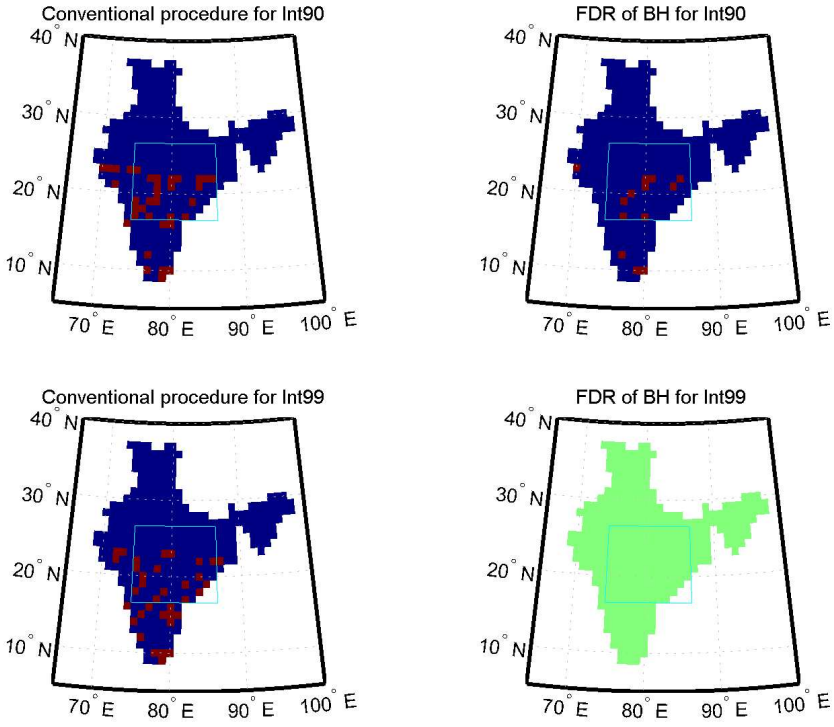


Figure 5.4: Same as Figure 5.3 but for grids with negative trends

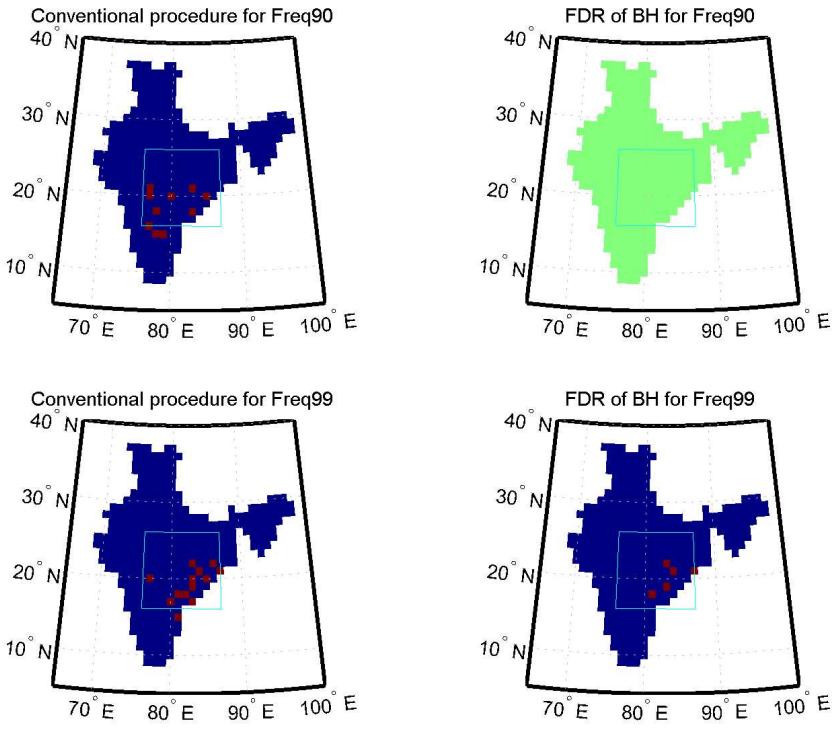


Figure 5.5: Spatial locations of the grids deemed to have positive significant trends in frequency of exceedance using the conventional and the BH procedures. Significance computed only for grids in the "core monsoon" region

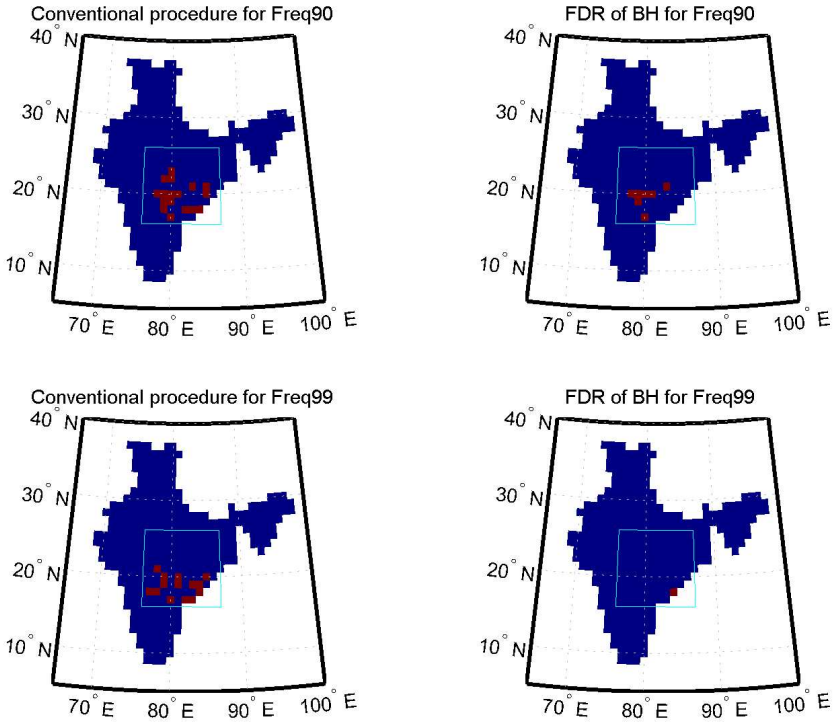


Figure 5.6: Same as Figure 5.5 but for negative trend

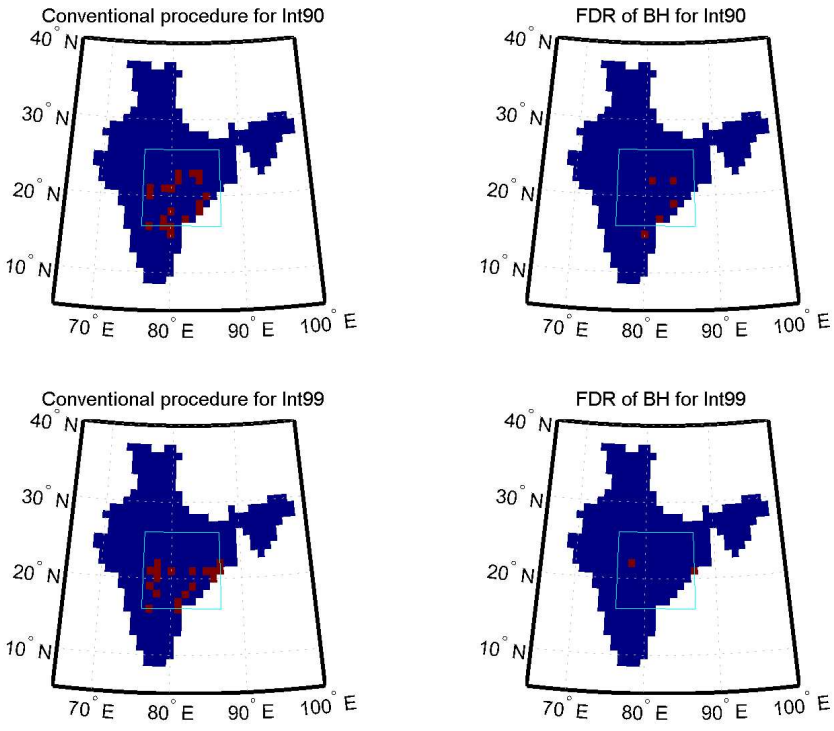


Figure 5.7: Spatial locations of the grids deemed to have positive significant trends in Intensity using the conventional and the BH procedures. Significance computed only for grids in the "core monsoon" region

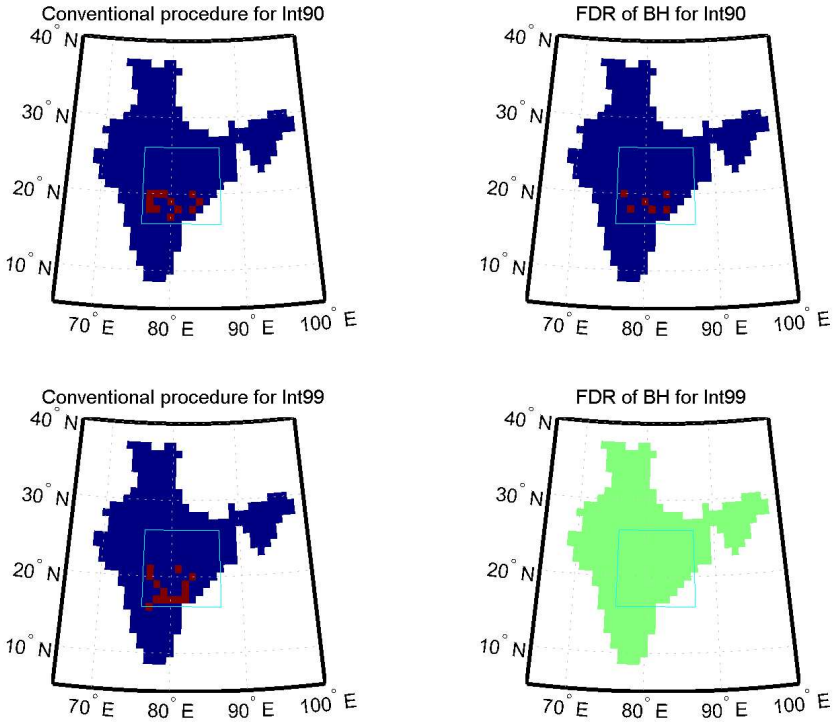


Figure 5.8: Same as Figure 5.7 but for negative trends

5.2.4 Spatial patterns in trends in frequency and intensity

From Figure 5.1 and 5.2, it is evident that, for frequency of extreme rainfall (at both thresholds), it is in the core monsoon region that trends (in both directions) are concentrated, and this is true across both FPR and FDR control. In the case of positive trends, it appears that change is more evident at the higher threshold, where it is more difficult to detect and attribute change (due to a paucity of data in the tail of the distribution). Further, an interesting feature in the figures is the concentration of positive trends at the edges of the monsoon region (and along the East coast) and of negative trends in the more Central regions, away from the coast.

When we turn to intensity, however, the first feature observed in frequency is stronger still i.e. the core monsoon region is quite clearly the area of high concentration of grids with change (of both directions). However, negative trends appear only at the lower threshold, while positive trends are evident at both thresholds. There is thus, evidence for positive trends in the core monsoon region (and elsewhere) and weaker evidence for decreasing trends in much of India. However, the second feature observed in frequency is not true, and most of the change is evident inland, and this is true using both measures of significance (FPR and FDR). Thus, the picture that emerges is one of increasing frequency of intense rainfall at the coasts and of increasing intensity of extreme rainfall in the interior, and especially over the core monsoon region.

We now turn to compare our results with prior work on this area. Goswami *et al.* [2006] find increasing intensity and frequency of extreme rainfall in the core monsoon region (using a fixed threshold and pooling data), while Krishnamurthy *et al.* [2009] perform a spatio-temporal analysis similar to the one in this analysis, and find weaker support for increasing frequency and slightly stronger support for increasing intensity. This study in fact rejects a field significance test based on the smaller core monsoon region, and in order to formally compare results with that study, we carry out a separate FDR procedure within the core

monsoon region.

While both of the previous papers used data for a shorter time period (1951-2004), the current very long record allows a better discrimination of trends especially when testing for trends in both directions, as in Krishnamurthy *et al.* [2009] (since the sample sizes in the shorter data set were too small for any reliable conclusions to be drawn). Thus, while earlier literature, using gridded data for a shorter time period, found mixed evidence for trends at a large scale, the direction, and locations, of the trends in the present analysis is largely in conformity with the previous results in Krishnamurthy *et al.* [2009].

Finally, Rajeevan *et al.* [2008] analyze, in a similar framework as Goswami *et al.* [2006], but with the same data set as we use, the issue of trends over the Central Indian region (the "core monsoon" region), and report similar results. Further, they carry out a preliminary analysis to link variations in the very-heavy rainfall events to tropical Indian ocean SST as well as ENSO and the IOD.

In order to render our results more comparable to prior studies, we also compute the significance of our results by restricting our analysis to the core monsoon region[5].The results of applying simple FPR procedures and the BH FDR procedure to grids in the "core monsoon" region are indicated in Figures 5.5 to 5.8, and in Table 5.2.

It is evident from these that there does not appear to be any significant differences in the number of grids with positive and negative trends, over both frequency and intensity. Consistent with the prior national-level analysis, at the higher threshold, intensity has no negative trends but also very few positive trends, indicating an overall lack of significant change. Further, the overall picture which emerges is one of weakly increasing trends at the higher threshold, and very uncertain, but possibly mostly negative, changes at the lower

[5]In other words, we use only grids in the core monsoon region for computation of significance using the BH approach. We do not make any claims as to the statistical validity of the approach and merely content ourselves with pointing out that this heuristic approach will enable us to understand if any significant changes are seen (relative to our full analysis). Evidence of lack of such changes must, we argue, serve to validate our all India analysis.

	Positive trends		Negative trends	
	FPR	**FDR**	**FPR**	**FDR**
frequency (90^{th} **%tile**)	10	0	17	7
frequency (99^{th} **%tile**)	13	5	14	1
Intensity (90^{th} **%tile**)	19	5	12	6
Intensity (99^{th} **%tile**)	18	2	14	0

Table 5.2: Same as table 1 but for grids in the core monsoon region (lat: 16.5 - 26.5, long: 76.5-87.5)

threshold. Our conclusions therefore reinforce those in Krishnamurthy *et al.* [2009], who found very similar results using a much shorter data set.

5.3 Spatio-temporal variability in extremes

We remind the reader that (monotonic) trends (investigated above) are only *one* aspect in which interest centers, as indicated in Section 5.1. In this section, we turn to analyzing the issues of cyclical behavior of extremes. In particular, we address the following:

1. What are the dominant spatial and temporal modes of behavior of extremes?

2. Do extremes exhibit cyclical behavior?

Diagnosis of important modes of large scale fields require an assessment of coherent spatio-temporal variation, for which the frequency-domain MTM-SVD approach is best suited.

5.3.1 MTM-SVD methodology

Identification of the important (low and high-frequency) modes of large-scale climate signals require an assessment of coherence in spatio-temporal climate variables. Most time-domain decomposition approaches are able to isolate broad patterns of variability, while being unsuited for isolating the narrow-band frequencies of most interest in climate data sets. The

principal frequency-domain approach for multivariate data sets, the multi-channel SSA (Ghil *et al.* [2002]), has severe limitations on both frequency bands over which the temporal structure maybe reconstructed as well as dimensional limitations for large data sets (see for instance Mann and Park [1999]). See Section 5.A for a more detailed explanation of the methodology.

5.3.2 Signal detection

The frequency domain MTM-SVD methodology involves the decomposition of a set of time series into its orthogonal spectral and spatial components by means of a singular value decomposition. In this manner, spatially coherent variability is sought within all distinct frequency bands that are resolvable in the dataset. Assuming that climate signals are relatively narrowband in character, a spectral domain approach enhances signal detectability. An important assumption regarding the data is that the power spectrum of the noise varies smoothly, so that it may be seen as roughly constant over the narrow bandwidth of the spectral estimates.

In the present analysis, the LFV spectrum is used to identify "significant" frequencies, and spatial reconstruction (see below) of the chosen field is carried out at these frequencies.

5.4 Spatio-temporal coherence of extremes

The question we attempt now to answer is: are there any common modes (frequencies) in the extremes (we use the same two measures, frequency and intensity, defined above). We first attempt to evaluate modes in each measure separately, and then analyze them together and identify the dominant frequencies.

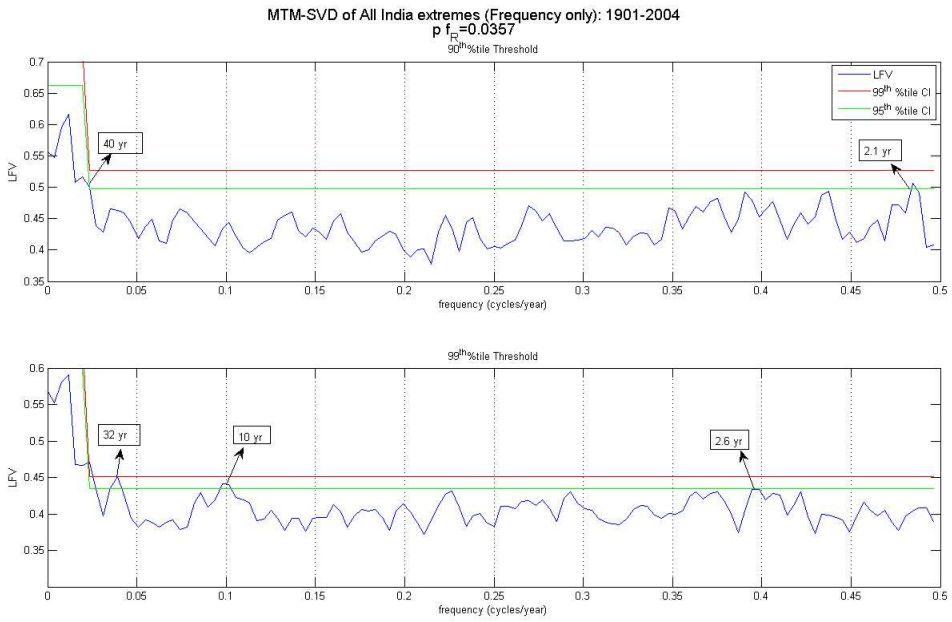


Figure 5.9: MTM-SVD of frequency of exceedance at two thresholds

Frequency

We first note that secular trends which are spatio-temporally coherent are seen as a spike at zero frequency. Given that we cannot resolve frequencies which are less than 0.0357 cpy apart, the approximately 40-year (and above) cycle, seen at both thresholds, is indistinguishable from a trend. In light of the known lack of robustness of most trends tests to long-term cycles, in subsequent analysis, we do not distinguish between very long-term cycles and secular trends and call all such features "trends" .

Other features of interest in the figure include the decadal variation (~10 year cycle) in the 99^{th}percentile threshold, as well as higher frequency variations (2.1 and 2.6 year) at both thresholds. Another feature is the absence of an ENSO-like (3-5 year) band, suggesting that the variations in frequency of exceedance, at a large enough scale at least, are not directly dependent on ENSO, and maybe modulated by other large scale movements.

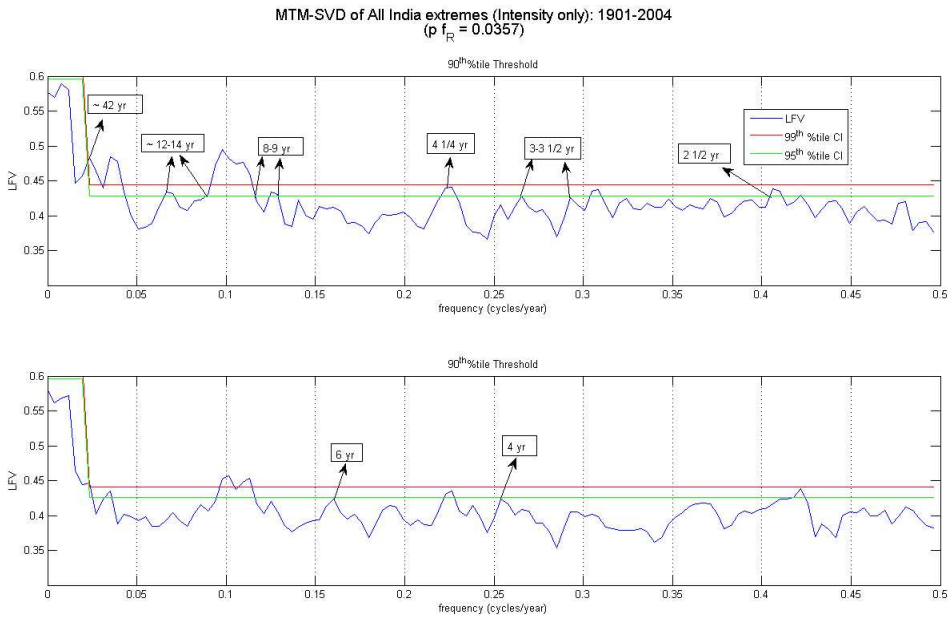


Figure 5.10: Same as Figure 5.9 but for Intensity

In sum, while trends and high-frequency variations are evident at both thresholds, decadal and ENSO-band variations are prominent only at the higher threshold.

Intensity

In the case of intensity, there appears to be much more cyclic behavior; for instance, there are, in addition to trend-like behavior and decadal and sub-decadal variations (12-14 year and 8-9 year), also two bands, at each threshold, which appears to be ENSO-like ($4^1/_4$ and $3 - 3^1/_2$ yr; 6 yr, 4 yr) and a final high frequency variation, as in frequency.

5.4.1 Joint MTM-SVD of frequency and intensity

From the previous sections analysis of trends in frequency and intensity separately, it is evident that there is no clear and coherent spatio-temporal picture of both. In order to analyze the variations in both collectively, we now carry out a joint MTM-SVD analysis

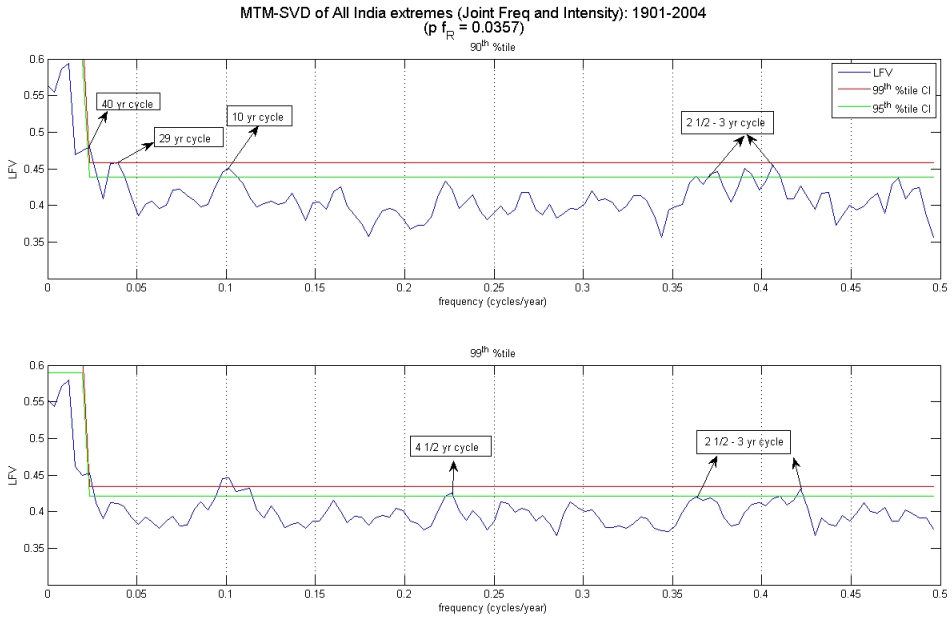


Figure 5.11: Joint MTM-SVD of frequency and intensity

of frequency and intensity, graphed in Figure 5.11. Consider first the lower threshold; we note that the results indicate existence of a trend (or long-term cycle) in both frequency and intensity, further, we also observe the decadal and multi-decadal variations, which were observed in both individually. However, at the higher threshold, in addition to these features, we also observe an ENSO-like band at $4^1/_2$ years, as in intensity.

In summary, from an individual as well as collective analysis of extreme events, we note that there are trend-like features, as well decadal and multi-decadal, high frequency features, and in some cases, ENSO-like bands, of large-scale variation in extremes. It is worth reiterating that, in the MTM-SVD methodology, "significant" features are shared by much of the spatial domain, and therefore, correspond to a signal at a relatively large spatial scale.

5.5 Conclusions

A more comprehensive analysis of extremes of rainfall was carried out than is common in the literature. In particular, there were two specific objectives:

1. To understand if secular trends identified in prior analyses, using a shorter (53 year) dataset and not accounting for multiple testing, are robust or an artifact of the data and methods used

2. To understand what additional features of extremes, in particular, quasi-oscillatory behavior, may be detected over a large region?

As regards question 1 above, it was shown that the overall results were consistent with the analysis in Chapter 1 (Krishnamurthy *et al.* [2009]). More detailed spatial structure was shown, using methods suited to identifying specific grids which exhibited significant trends. Thus, the overall picture which emerges is one of weakly increasing trends at the higher threshold, with very uncertain (even negative) trends at the lower threshold.

On question 2 above, the features identified vary with the metric for extreme used. For the case of frequency-of-exceedance, large scale coherence exists at the $ENSO$ and the decadal time periods. This lends some support to conjectures in the literature (Rajeevan *et al.* [2008], for instance) regarding large-scale spatio-temporal connections with frequency of extreme rainfall events. On the other hand, for intensity of extreme rainfall, there is no clearly discernible signal, with ENSO and decadal and sub-decadal features appearing significant.

The context for understanding these results is the substantial importance that changes in extremes have on society, in particular for agriculture and infrastructural decisions, and the issue of predictability of extremes from (or at least association of extremes to) large-scale, observable physical phenomena. Under climate change scenarios, ENSO activity can intensify (Meehl *et al.* [2007]); when coupled with the increased atmospheric water retention

capacity, there is a possibility of increased extreme rainfall events, among other extremes (Smith *et al.* [2009]). Given the concerns regarding the increases in damages caused by such events, decisions on both mitigation (see Stern [2007]) and adaptation are likely significantly impacted by such putative increases in rainfall extremes.

Analyses such as those carried out here can provide an understanding of the large-scale connections between societally-relevant outcomes (extremes) and physical phenomena (ENSO etc), providing a basis for prediction, on the one hand, and help inform mitigation, on the other.

Bibliography

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.

M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. Advanced spectral methods for climatic time series. *Rev. Geophys.*, 40, 2002.

B.N. Goswami, V. Venugopal, D. Sengupta, M.S. Madhusoodanan, and Prince K. Xavier. Increasing trends of extreme rain events over india in a warming environment. *Science*, 314:1442–1444, December 2006.

P.Y. Groisman, R.W. Knight, D.R. Easterling, T.R. Karl, G.C. Hegerl, and V.N. Razuvaev. Trends in intense precipitation in the climate record. *Journal of Climate*, 18(9):1326–1350, 2005.

Edward L. Korn, James F. Troendle, Lisa M. McShane, and Richard Simon. Controlling the number of false discoveries: application to high-dimensional genomic data. *Journal of Statistical Planning and Inference*, 124(2):379 – 398, 2004.

Chandra Kiran B. Krishnamurthy, Upmanu Lall, and Hyun-Han Kwon. Changing frequency and intensity of rainfall extremes over india from 1951 to 2003. *Journal of Climate*, 22:4737–4746, 2009.

M.E. Mann and J. Park. Oscillatory spatiotemporal signal detection in climate studies: A multiple-taper spectral domain approach. *Advances in Geophysics*, 41:1–131, 1999.

G.A. Meehl, T.F. Stocker, W.D. Collins, AT Friedlingstein, A.T. Gaye, JM Gregory, A. Kitoh, R. Knutti, JM Murphy, A. Noda, et al. Global climate projections. 2007.

Mark New, Bruce Hewitson, David B. Stephenson, Alois Tsiga, Andries Kruger, Atanasio Manhique, Bernard Gomez, Caio A. S. Coelho, Dorcas Ntiki Masisi, Elina Kululanga, Ernest Mbambalala, Francis Adesina, Hemed Saleh, Joseph Kanyanga, Juliana Adosi, Lebohang Bulane, Lubega Fortunata, Marshall L. Mdoka, and Robert Lajoie. Evidence of trends in daily climate extremes over southern and west africa. *J. Geophys. Res.*, 111, |2006|.

Thomas C. Peterson, Xuebin Zhang, Manola Brunet-India, and Jorge Luis Vázquez-Aguirre. Changes in north american extremes derived from daily weather data. *J. Geophys. Res.*, 113, |2008|.

Balaji Rajagopalan, Michael E. Mann, and Upmanu Lall. A multivariate frequency-domain approach to long-lead climatic forecasting. *Weather and Forecasting*, 13(1):58–74, 1998.

M. Rajeevan, Jyoti Bhate, and A. K. Jaswal. Analysis of variability and trends of extreme rainfall events over india using 104 years of gridded daily rainfall data. *Geophys. Res. Lett.*, 35, 2008.

J Schmidli and C Frei. Trends of heavy precipitation and wet and dry spells in switzerland during the 20th century. *International Journal of Climatology*, 25(6):753–772, |2005|.

J.B. Smith, S.H. Schneider, M. Oppenheimer, G.W. Yohe, W. Hare, M.D. Mastrandrea, A. Patwardhan, I. Burton, J. Corfee-Morlot, C.H.D. Magadza, et al. Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC)"reasons for concern". *Proceedings of the National Academy of Sciences*, 106(11):4133, 2009.

N.H. Stern. *The economics of climate change: the Stern review*. Cambridge Univ Pr, 2007.

V Ventura, CJ Paciorek, and JS Risbey. Controlling the proportion of falsely rejected

hypotheses when conducting multiple tests with climatological data. *Journal of Climate*, 17(22):4343–4356, 2004.

D. S. Wilks. On "field significance" and the false discovery rate. *Journal of Applied Meteorology and Climatology*, 45(9):1181–1189, 2006.

Panmao Zhai, Xuebin Zhang, Hui Wan, and Xiaohua Pan. Trends in total precipitation and frequency of daily precipitation extremes over china. *Journal of Climate*, 18(7):1096–1108, |2005|.

Appendix 5.A MTM-SVD Methodology

We will here provide a brief overview of the MTM-SVD methodology (see Mann and Park [1999] for a detailed analysis and comparison with other multi-variate decomposition methodologies), with a focus on application to the analysis of extremes of rainfall (the description below draws from Rajagopalan *et al.* [1998] and from Mann and Park [1999]).

In more detail, each of the M time-series, of length N, is first standardized by converting it to an ''anomaly'' series i.e. removing the series mean, and dividing by the standard deviation. For a time-series $y = (y_1, \ldots, y_N)$, at location $m = 1, \ldots, M$, using a set of K orthonormal data tapers (see below), K tapered Fourier transforms of the data at a frequency f are obtained i.e.

$$Y_k^m(f) = \sum_{t=1}^{N} w_t^k y_t e^{2\pi i t \Delta t} \tag{5.6}$$

with Δt being the sampling interval (1 year, in our case), and $\{w_t^k\}_{t=1}^N$ is the k^{th} member, $k = 1, \ldots, K$, of an orthogonal sequence of tapers, known as *Slepian tapers*. The Slepian tapers provide statistically independent spectra estimates, $Y_k^m(f), k = 1, \ldots, K$, leading to multiple degrees of freedom within a given narrow frequency band, and furnishes the primary motivation for the use of the multitaper decomposition. Each eigentaper is orthogonal and represents independent information in a frequency band of half-bandwidth of pf_R about a given frequency, f, where $f_R = (N\Delta t)^{-1}$ is the Rayleigh frequency (the minimum resolvable frequency range for the time series). Since a larger p averages over a greater bandwidth, the choice between K and p represents the fundamental trade-off between spectral resolution and degrees of freedom/variance. Values of 2 and 3 for K and p respectively, are seen as typical for climate data sets and are used in this analysis.

An $M \times K$ matrix $A(f)$ is then formed from the K fourier transforms, as below, for each frequency f,

$$A(f) = \begin{bmatrix} Y_1^1 & y_2^1 & \dots & Y_K^1 \\ Y_1^2 & y_2^2 & \dots & Y_K^2 \\ \dots & \dots & \dots & \dots \\ Y_1^M & y_2^M & \dots & Y_K^M \end{bmatrix} \tag{5.7}$$

Note that the rows correspond to spatial locations (grids, in our case) and columns to the tapers. Isolation of spatially-coherent narrowband processes is through a complex singular value decomposition (SVD) of the above matrix

$$A(f) = \sum_{k=1}^{K} \lambda_k(f) \mathbf{u_k}(f)^* \mathbf{v_k}(f) \tag{5.8}$$

into K orthonormal complex $M-vectors$ $\mathbf{u_k}$ (the left eigenvector), representing complex spatial EOF's and K orthonormal $K-vectors$ $\mathbf{v_k}$(the right eigenvectors), termed conventionally "spectral EOF's". The singular value $\lambda_k(f)$ scales the amplitude of the k^{th}mode in this local eigen-decomposition (with the K singular values being ordered $\lambda_1(f) \geq \lambda_2(f) \geq \dots \geq \lambda_K(f)$), with $\lambda_k(f)^2$ being deemed "eigen values". The normalized "principal eigen value", $\lambda_1^2(f)/\sum_{j=1}^{K} \lambda_j^2(f)$, referred to as the "local fractional variance spectrum" (LFV) , provides a signal detection parameter that is *local* in the frequency domain. The normalized principal eigenvalue should, assuming no more than one signal is present within the narrow bandwidth of spectral estimation, stand out distinctly above what would be expected from an appropriate noise model. The confidence levels are based on a locally-white-noise null, are constant outside the secular band, and are computed using a bootstrap approach.

CHAPTER 6

Structural Properties of Groundwater Models[1]

[1] Joint work with Tim Huh. A small part of this chapter, Section 6.6, is being revised for resubmission at the *Naval Research Logistics*.

6.1 Introduction

There is a substantial economic literature on groundwater management, in a single-cell aquifer, by a single user (beginning with Burt [1964], described in Koundouri [2004]). In this chapter, we consider the problem of managing groundwater under random recharge in a single cell aquifer. There are two main modeling paradigms used in the literature: the first based on optimal control of the continuous-time model, e.g., Brown and Deacon [1972], Gisser and Sanchez [1980], Tsur and Graham-Tomasi [1991], Hellegers *et al.* [2001], and Tsur and Zemel [2004], Zeitouni [2004], Roseta-Palma and Xepapadeas [2004], Rubio and Casino [2001] and the second based on dynamic programming formulation of the discrete-time model, e.g., Burt [1964, 1966, 1967, 1970], Provencher and Burt [1993] and Knapp and Olson [1995, 1996].

Most continuous time models, due to the specific assumptions made, yield explicit solutions. For instance, in the linear case, there are deterministic (Aggarwal and Narayan [2004]) or stochastic (Zeitouni [2004]) "bang-bang" solutions while in the linear-quadratic case (Roseta-Palma and Xepapadeas [2004]) there are analytical solutions possible. There have been very few continuous time models wherein structural properties have been derived from relatively few assumptions on the primitives. While several discrete-time dynamic programming formulations have been proposed in the literature, few structural properties for this problem have been proven. In fact, it is only Knapp and Olson [1995] who provide a first proof of the relatively straightforward property that extraction is increasing in current period stock.

Further, almost all of these formulations, both discrete- and continuous-time, make strong assumptions regarding the cost of extraction. In particular, two assumptions are made, usually implicitly: (a)marginal cost of pumping is independent of the quantity pumped (Chakravorty and Umetsu [2003] explicitly state this assumption) and (b) cost

of pumping depends only on the beginning of period stock. Both of these are assumptions, with little empirical support even for unconfined aquifers (another assumption implicit in many studies). Yet, even with these assumptions, most structural properties of interest (see below) have not been proved.

Formally, we consider here the canonical discrete-time stochastic version of the model (see §6.2.1 for details). That this setting, for a wide variety of renewable resource problems, is similar to the standard stochastic one-sector neo-classical growth model has been documented extensively (beginning at least with Mendelssohn and Sobel [1980]). Many resource problems (such as those on resource extinction, see Olson and Roy [2000] and references therein or in Mitra and Roy [2006]) are dealt with in the setting of the stochastic growth model. However, given that the model can be solved independently in a dynamic programming framework, we do not emphasize such a linkage here. In such models, structural properties, in particular Properties (a),(c), (d) and (f) (in an infinite horizon model) below are of interest.

- **Property** (a): The optimal withdrawal quantity in period t, $w_t^*(x)$, is increasing in x,
- **Property** (b): The optimal withdrawal quantity in period t, $w_t^*(x)$ (which is a function of the groundwater stock x at the start of the period), is the maximizer of a concave function of w,
- **Property** (c): $w_t^*(x)$ is nondecreasing in x,
- **Property** (d): $x - w_t^*(x)$ is nondecreasing in x, and
- **Property** (e): $w_t^*(x)$ is nondecreasing in t, where the periods are indexed forward in time
- **Property** (f): The Markov chain generated by the optimal policy, $w_t^*(x)$, $\{X_t\}$, converges to a unique, stationary distribution.

We note first that in the case of the groundwater models, many of the results in Mendelssohn and Sobel [1980] are not directly applicable (we indicate, after each result, if it follows from theirs or any other framework). Further, the question regarding the concavity of the value function has been answered in the negative in Knapp and Olson [1995]. Models used in the literature are replete with assumptions regarding smoothness, in particular, (joint) concavity of the objective function and convexity of the relevant spaces. Such assumptions appear to be an artifact of the sufficient conditions for obtaining the structural properties listed above than arising from any underlying characteristic of the natural-economic system.

Knapp and Olson [1995, 1996] move away from such smoothness conditions and work in a lattice-theoretic framework, which we also adopt. However, their formulation of the problem side-steps the issue of uncertainty (indeed, their model is very similar to a deterministic model). In their set up, uncertainty (in surface water flows) is resolved *prior* to the farmer making the extraction decision. In such a set up, the farmer can directly control the succeeding periods stock, as a result of which they work directly with the next periods stock. Thus, proof of monotonicity of next periods stock in current period stock is sufficient for them to apply to theorems on convergence of Monotonic Markov Chains.

On the other hand, in this set up, *uncertainty has no role to play*, from the decision perspective. The model we use, on the other hand, involves uncertainty being resolved after the farmer has made extraction decisions, and corresponds better to a real world scenario in a developing nation wherein farmers make decisions on extraction prior to recharge (rainfall) occurring. In this set up, the inability to directly control the subsequent periods stock leads to uncertainty playing a central role in the farmers extraction decisions. Further, monotonicity of extraction and reinvestment are not identical, as we indicate, even for the model studied in Knapp and Olson [1995, 1996].

The major objectives of this chapter are to provide a unified treatment of the structural

properties of the dynamic programming problem of groundwater extraction. The unification is both in terms of methods of proof used (we rely for the most part on lattice-theoretic methods) as well as on using very general, realistic cost functions for extraction. We also move completely away from smoothness assumptions, in particular abandoning all assumptions of concavity of the objective functions (unlike for instance in Mendelssohn and Sobel [1980]), relying only on the monotonicity and supermodularity of the objective function.

Finally, as an adjunct of the smoothness assumptions mentioned, the Markov Chain generated by the optimal policy is monotone. This allows (as in Knapp and Olson [1995, 1996]) a direct application of standard theorems on convergence of Monotone Markov processes (as popularized in economics in Stokey and Lucas [1989] and Hopenhayn and Prescott [1992]) to establish "global stability" of the resulting Markov Chain. Even here, however, compactness of the state space (implying boundedness of the shocks) is an essential assumption of the method of proof. In this setting, we illustrate the use of slightly more probabilistic arguments which guarantee the existence of an invariant distribution even in cases where the Markov Chain is *not monotonic*. The assumptions required for the use of this method is more bengin than even for the monotone case using more conventional methods in economics. In other words, with weaker assumptions, it is possible to prove *stability* of the Markov Chain generated by the decision process, even in cases where the Markov Chain is *not* monotone. These methods are applicable to a wide variety of natural resource problems, and are of independent interest.

6.2 Structure of the Models considered

6.2.1 Model Setup

We describe a discrete-time stochastic groundwater management model, based on the classical paper by Burt [1964], which is standard in the literature. We consider a finite-

horizon problem where T is the planning horizon, and periods are indexed forward by $t \in \{1, 2, \ldots, T\}$. Let $x_t \geq 0$ denote the groundwater stock[2] at the beginning of period t. The manager decides the withdrawal quantity $w_t \geq 0$. Then a non-negative random variable corresponding to the recharge to the groundwater stock, R_t, is realized. We suppose that the $\{R_t\}$ are independent and identically distributed, but indicate generalizations where appropriate and feasible. We assume that the state transition for the groundwater stock level is given by[3]:

$$x_{t+1} = \tilde{X}(x_t, w_t) = x_t - w_t + R_t \tag{6.1}$$

While it is conventional to use a $\max(z, 0)$ operator to signify that the $x_t$is taken non-negative, we do not do so at this stage since it is implicit in our analysis and further, for the cost function in (6.6) below, such a restriction is not even essential for subsequent results.

The single-period benefit (or reward) is

$$G(x, w) = B(w) - C(x, w) \tag{6.2}$$

where $B(w)$, concave and increasing, is the net benefit of withdrawing w units of water before deducting the pumping cost $C(x, w)$ (which we will address shortly below). The objective is to maximize the present value of the benefit

[2]Following the Resource Economics literature (see papers cited above), we work with the total stock of water instead of the *lift*, which is more conventional in the engineering literature. However, as remarked in Worthington *et al.* [1985](pp 235), since lift increases as groundwater decreases, both are related via a monotonic function and therefore, one may work with either of them.

[3]The use of such a simplified balance equation, as remarked in Worthington *et al.* [1985](pp 232-233) is a gross over simplification. Taken literally, this equation implies an "instantaneous" capture of all recharge by any pumping activity. However, this oversimplification can be remedied first by introducing relevant coefficients on recharge such that only a fraction of the recharge is captured. Second, in the case of many developing nations where discharge, due to pumping, is much larger than recharge, such as assumption is less of an oversimplification, since a large part of recharge is very likely captured within the region of pumping. Finally, more detailed and accurate equations of motion of the stock may be used where such a relationship is available.

$$\sum_{t=1}^{T} \delta^t G(x_t, w_t) \tag{6.3}$$

where $\delta \in (0, 1]$ is the discount factor. Then, the value function of dynamic programming is given by the following recursion:

$$V_t(x_t) = \max_{w_t \geq 0} \{G(x_t, w_t) + \delta E[V_{t+1}(x_t - w_t + R_t)]\} \tag{6.4}$$

where $V_{T+1}(x_{T+1}) = 0$[4].

We note that, unlike in much of the Economic literature on the one-sector stochastic growth model, wherein focus is on the infinite horizon problem, we begin with a finite horizon set up and indicate extensions to an infinite horizon. The reason for this two step approach is the added understanding provided in the finite horizon set up, especially with regard to such intuitive questions as the behavior of the optimal policy with a lengthening horizon. This approach also makes transparent which type of assumptions regarding the state space and the benefit functions maybe relaxed. We turn next to characterizing the cost functions we use, and comparing them with those used in the literature.

6.2.2 Cost Functions

Pumping cost functions used in the literature mostly make the assumption of constant volume pumped as well as of constant (relative to w_t) marginal cost of pumping i.e. $\frac{\partial C(x_t, w_t)}{\partial w_t} = c(x_t)$. Within this broad framework, with the exception of Burness and Brill [2001]; Worthington *et al.* [1985], most papers (Aggarwal and Narayan [2004]; Provencher and Burt [1993]; Roseta-Palma and Xepapadeas [2004]; Rubio and Castro [1996]; Tsur and

[4]The use of the max operator instead of sup indicates an assumption of finiteness of $G(x, w)$, since such an assumption is sufficient for the right hand side of (6.4) to be finite. The above usage is merely illustrative, and in specific models below, we outline the assumptions we make to ensure that the use of the max operator is valid in this setting.

Graham-Tomasi [1991]; Zeitouni [2004]) use a cost function of the form

$$C(x_t,w_t) = c(x_t)w_t \tag{6.5}$$

with $c(x_t)$ (generally called the marginal cost function) either linearly decreasing in x_t or decreasing and convex, with the most common functions being $c(x_t) = (a - bx_t)$ or $\frac{a}{bx_t}$ (except for Roseta-Palma and Xepapadeas [2004] who use $c(x_t, w_t) = aw_t$, a stock independent linear cost). Burness and Brill [2001] use a slightly different formulation for $c(x)$ while maintaining the separability outlined above. Their marginal cost function (in our terminology) is of the form $c(x) = \frac{c_0(a-x)}{bc(x-d)}$ with $a, c, b, d, c_0 > 0$ and $x > d$. This function also has all the usual properties $(c(x)_x < 0, c(x)_{xx} > 0)$ that the simpler functions above do. Finally, Worthington *et al.* [1985], who also use a separable cost function, use an empirical form of the marginal cost function, estimated (interpolated) from data. Their function is decreasing but, due to being a combination of different order polynomials for different regions, has $c(x)_{xx} \lesseqgtr 0$ depending upon x.

In general, all cost functions used make the following two assumptions

1. Marginal cost of pumping does not depend upon the quantity pumped i.e. $\frac{\partial C(x_t, w_t)}{\partial w_t} = c(x)$

2. Cost of pumping does not vary within a given season i.e. drawdown does not vary within a given season. This leads to specifications in which marginal costs are based entirely on beginning-of-period water level depth. This assumption is completely implicit in almost all studies, discrete- and continuous-time, and is made clear only in Worthington *et al.* [1985] (pp 235-237) wherein (in their Table 1, pp 236) they estimate an *annual incremental drawdown* given only the level of water table at the *beginning* of the season, *independent* of the quantity of water pumped.

The first characteristic is a modeling assumption and is not always in accord with hydrologic

facts (see below). The second characteristic above is reasonable if the length of a period is sufficiently small and decisions are made frequently, in which case continuous-time models are more appropriate. In sum, in all of the literature, marginal costs of pumping do not vary within season, and are independent of the quantity pumped during the period. The implications are that, for the marginal cost of pumping in any *given period*, the only factor of importance is the groundwater stock at the beginning of the period; the amount of abstraction during the period does not matter. Put differently, the implications are rather striking: the marginal costs of pumping, given a beginning stock, are the same, irrespective of the quantity extracted.

With this cost function, structural properties such as the concavity of the optimization problem in each period have not been established[5]. As noted in Section 6.1, there have been relatively few efforts focused on establishing structural properties of discrete-time groundwater models in the literature, and many of the properties still have not been shown to hold, even in a setting with very simple (and unrealistic) cost functions. In our analysis, we generalize this formulation for cost to reflect more realistic conditions on marginal cost and study in detail the structural properties of the resulting dynamic programs. In particular, we seek to address two important questions:

1. Under what conditions, for the given objective function(s), do intuitive properties (such as monotonicity of withdrawal, next periods stock in current period stock and existence of a stationary distribution for the stochastic stock) hold?

2. Which of these properties hold under a setting where risk aversion is introduced?

We carry out two generalizations of the cost function. In the first,

[5] Provencher and Burt [1993][pp 146]show the concavity of the value function, but their result crucially depends on (a) the separability of $C(x, w)$ (b) the fact that the per-unit pumping cost is independent of the quantity, w_t, of groundwater extracted and (c) an *unsubstantiated* claim of concavity of the objective function. See footnote (9)

$$C(x,w) = \int_{z=x-w}^{x} \gamma(z)dz \tag{6.6}$$

where γ is assumed to be non negative, decreasing and convex. This form of the cost function takes account of changes in the groundwater stock that occur during the pumping period. Further, the marginal cost of pumping depends, in general, on the quantity of water pumped. We stress that the conventional approach, in equation (6.5), with the implication that marginal cost of pumping is independent of the quantity pumped, is an assumption (recognized explicitly in Chakravorty and Umetsu [2003](pp 5)), and has little empirical or theoretical support.

In the second, we generalize the cost function in a different direction. In many cases, and especially in developing nations (such as India) where landholding sizes tend to be far smaller, many wells are in close proximity, in which case, many assumptions of the "bath tub" aquifer are not very tenable, especially when transmissivity is limited (which amplifies the impact, on a given user, of pumping by other users). There has been an increasing body of literature (see for instance Brozovic *et al.* [2002]; Chakravorty and Umetsu [2003]; Zeitouni and Dinar [1997], among others) dealing with a variety of issues, especially on the game-theoretic aspects of managing such common pool resources. The main issue involved is the formation of "cones of depression" around individual wells, leading to increasing extraction costs. These costs depend, to a large extent, on a variety of physical parameters such as transmissivity and storativity (see for instance Athanassoglou *et al.* [2011]).

On the other hand, even in the case of a single user, the assumptions made in almost all models regarding the aquifer are unlikely to hold for any given aquifer. For instance, most continuous time models, following Gisser and Sanchez [1980], assume an essentially bottomless aquifer. Further, most analyses (exceptions include Worthington *et al.* [1985]) assume implicitly or explicitly an unconfined aquifer. However, for confined aquifers, ongoing pumping from a well induces a localized "cone of depression" around the well (see

for instance Siegfried [2004] (Figure 4.4, pp 61)). These dynamics lead to increases in the per-unit cost of extraction by increasing the effective lift[6].

We seek a simplified framework to account for the increased marginal cost of pumping as a result of such localized cones of depression. We use the finite difference cell approach used in Siegfried [2004] (pp 52-53). The method effectively corrects the drawdown by adding a term linear in the extraction. Abstracting away from parametrization, we translate this into our notation as follows:

$$C(x, w) = c\,(\bar{x} - x + aw)\,w \tag{6.7}$$

where the additive term aw captures the "correction" to the drawdown (change in stock)[7].

6.3 Analysis: Finite Horizon

We consider the dynamic model of groundwater extraction using each of the three cost functions, separately, since many features, method of proof and conclusion differ. The approach taken in our analysis is somewhat different from that conventionally used in the Resource Economics (and Economic Growth) literature. We abandon, for the most part, *all assumptions regarding smoothness* (concavity, convexity and differentiability) of the objective function and feasible set, except insofar as these properties can be *proved*

[6]This issue, even with the confined aquifer assumption, has been remarked on before. For instance, Provencher and Burt, 1994 (pp 882) clearly recognize that groundwater levels do not adjust immediately after a local perturbation caused by pumping from the well, due to the fact that the equilibrating flows are both slow and subject to great variability. In particular, with heterogeneous aquifers and unequal pumping rates by farmers, they clearly recognized a not insubstantial lateral flow of groundwater, leading to uncertainty of future availability of water. While their discussion was in the context of a common property problem, the discussion is clearly applicable in the context of our modification to the cost function, even if one ignores the common property aspect, as we do in our work here.

[7]For subsequent analysis, we drop the coefficient a on w for minimizing the number of constants used. This coefficient maybe introduced as needed for empirical analysis, and its exclusion does not fundamentally impact our analysis but does lead to simpler algebraic manipulations.

from physically or economically reasonable primitives. We use lattice theoretic methods, developed in Topkis [1978], with a comprehensive account in Topkis [1998], and the detailed exposition of dynamic programming using lattice-theoretic methods developed in Heyman and Sobel [2003].

We enumerate below a list of assumptions, not all of which are used in all Theorems or with all cost functions. We also set up some notation for future reference. Let $X \subset R_+$be the state space, which is a lattice, $W : X \to L(X)$, with $L(X)$ the set of all sub-lattices of X, $\mathcal{C} = \{(x, w); x \in X, w \in W(x)\}$ be a sub-lattice of $\mathbb{R}^2_+$.

Definition 6.3.1. (Ascending Set-Valued functions) The set-valued function W is called *ascending in x on X* (or simply "ascending") if it is *increasing on X*. In this definition, if x_1 and x_2 are in X, then $W(x_1)$ and $W(x_2)$ are in $L(X)$. Therefore, if $x_1 < x_2$, $a \in W(x_1)$ and $b \in W(x_2)$ and W is ascending on X, then necessarily $a \wedge b \in W(x_1)$ and $a \vee b \in W(x_2)$ [8].

Definition 6.3.2. (Expanding Sets)The set $W(x)$ is called "expanding" if $x' < x \implies W(x) \subset W(x')$.

Definition 6.3.3. (Stochastically supermodular) $\tilde{X}(x, w)$, a random variable parametrized by (x, w), defined on $X \times X$, a lattice, $F_{x,w}$ its distribution function, is said to be "stochastically supermodular" in (x, w) if either of the following two conditions are satisfied:

1. [Topkis, 1998, pp 159] $\int_S dF_{x,w}(s)$ is supermodular in (x, w).
2. $\int h(s)dF_{x,w}(s)$ (as a deterministic function of (x, w)) is supermodular in (x, w), for all increasing and bounded functions h.

Definition 6.3.4. (Re-investment function)$a = x - w$ is defined to the "reinvestment function". It will be used in the proof of monotonicity in t of the value function and the

[8] For ex. if $x \in X \subset \mathbb{R}$, $W(x) = [-\infty, x]$ or $W(x) = [x, \infty]$ are ascending on X.

optimal policy function (where applicable). It is trivial to reformulate all of the statement and assumptions in terms of a instead of w. Let $A(x)$ be the equivalent of $W(x)$ for a.

Assumption 6.3.5. $\Pi(x, w)$ *is supermodular in* (x, w) *and increasing in* x

Assumption 6.3.6. $\Pi(x, a)$ *is supermodular in* (x, a) *and increasing in* x*.*

Assumption 6.3.7. $\tilde{X}(x_t, w_t)$ *is increasing in* x_t*, given* w_t, R_t*(or independent of* x_t*).*

Assumption 6.3.8. $\Pi(x, w) \geq 0$*,* $\forall x, w$

Assumption 6.3.9. Π *is finite on* $\mathcal{C}$

Assumption 6.3.10. $\exists B > -\infty$ *s.t.* $\Pi(x, w) \geq B$*,* $\forall x, w$ *i.e.* Π*is uniformly bounded below.*

Assumption 6.3.11. $\Pi(x, a) \geq 0$

Assumption 6.3.12. $W(x)$ *(*$A(x)$*) is ascending in* x *on* X*.*

Assumption 6.3.13. $W(x)$ *(*$A(x)$*) is expanding in* x*.*

Assumption 6.3.14. $W(x)$ *(*$A(x)$*) is compact.*

Assumption 6.3.15. $\tilde{X}(x_t, w_t)$ *is stochastically supermodular.*

Assumption 6.3.16. $\tilde{X}(x_t, a_t)$ *is stochastically supermodular*

We repeat here the dynamic programming recursion from (6.4), with

$$V_t(x_t) = \max \{J_t(x_t, w_t); x \in X, w \in W(x)\} \tag{6.8a}$$

$$J_t(x_t, w_t) = \Pi(x_t, w_t) + \delta \mathbb{E}[V_{t+1}(x_t - w_t + R_t)] \tag{6.8b}$$

6.4 Conventional Cost Function (equation(6.5))

6.4.1 Risk-Neutral decision making

Recall from equation(6.5) our cost function is:

$$C(x,w) = c(x)w = c\left(\bar{x} - x\right)w \tag{6.9}$$

Thus, the objective function (which we call Π for this section) is:

$$\Pi(x_t, w_t) = B(w_t) - c\left(\bar{x} - x\right)w \tag{6.10}$$

Two facts are evident from equation(6.4). First, Π is not jointly concave in (x,w)[9]. Second, Π is supermodular in (x,w). This is evident from the differential characterization i.e. a smooth (for instance $\mathcal{C}^2$) real valued function $f(u,v)$ on a lattice is supermodular in (u,v) iff $\dfrac{\partial^2 f(u,v)}{\partial u \partial v} \geq 0$. That this is unconditionally true is evident from the fact that $\dfrac{\partial^2 \Pi(x,w)}{\partial x \partial w} = c > 0$. Further, it is equally evident that $\Pi_x(x,w) = cw > 0$.

We make a further assumption before we state our main result for this section.

Assumption 6.4.1. *$J_t(x_t, .)$ is upper semi-continuous on $W(x_t)$ for each $t \in \{1, 2, \ldots, T\}$ and $V_{T+1}(x) = 0$*

We are now ready to state our main theorem

Theorem 6.4.2. *[Heyman and Sobel [2003],Corollary 8-5a, pp 383]Assumptions (6.3.5), (6.3.7),(6.3.8),(6.3.12),(6.3.13),(6.3.15),(6.4.1) imply that $V_t(x)$ (defined in (6.8a))is increas-*

[9]To see this, note that: $C_x = -c$, $C_{xx} = 0$, which leads to $\Pi_{xw} = -c$, $\Pi_{xx} = -C_{xx} = 0$, $\Pi_{ww} = -B^{''}$. For Π to be jointly concave in (x,w), the Hessian is required to be positive semi-definite, i.e. $B^{''}C_{xx} \geq (C_x)^2$ which in this case is equivalent to $c^2 \leq 0$, which is evidently false. Thus, Π does not satisfy a sufficient condition to be jointly concave. Of course, this condition is not necessary for concavity, but it is unlikely that Π will be jointly concave, since Π is otherwise very smooth.

ing in x and that $W_t^*(x) = argmax\{J_t(x_t, w_t); w_t \in W(x)\}$ is ascending in x on $\{x; x \in X, argmax\{J_t(x_t, w_t)\}$ is non-empty$\}$. Further there exists a least and a greatest element in the set $W_t^*(x)$ and these are both increasing in x.

Proof. We provide a detailed proof of the theorem, since many subsequent proofs will be based on, or refer to, this theorem. That $V_t(x)$ is increasing in x is a straightforward consequence of two facts, $V_{T+1}(x) = 0$ and $\Pi(x, w)$ is increasing in x, given w. These both facts can be used to set up a simple inductive argument, as below.

$V_{T+1}(x)$ is trivially increasing in x, and so the induction is true for $t = 0$. Similarly, $V_T = \max\{\Pi(x, w); w \in W(x)\}$ is increasing in x, since Π is increasing and $W(x)$ is ascending and expanding and so the hypothesis is true for $t = 1$. Let it be true for $t = k < T$ i.e. $V_k(x)$ is increasing in x. Consider $V_{k-1}(x) = \max_{w_t \geq 0}\{J_{k-1}(x_t, w_t\}$, where $J_{k-1}(x, w)$ is increasing in x, since both its terms are (the second term, $\mathbb{E}(V_k(x))$, is increasing by assumption above and the fact that integration is order preserving) and thus, so too is V_{k-1}. As an aside, observe that if Π is bounded so too is V_t.

We next prove that $\mathbb{E}\left[V_t\left(\tilde{X}(x_t, w_t)\right)\right]$ is supermodular in (x, w). We have proved that V_t is increasing and bounded, and have assumed that the distribution function of $\tilde{X}(x_t, w_t)$, $F_{x,w}$ is supermodular. Therefore, by Topkis [1998][Corollary 9.1(b), pp 160], $\mathbb{E}\left[V_t\left(\tilde{X}(x_t, w_t)\right)\right]$ is supermodular. Thus, we have that $J_t(x_t, w_t)$ in (6.8b) is supermodular (since the sum of two supermodular functions is supermodular[10]).

Since $W(x)$ is ascending and expanding in x, that $W_t^*(x) = argmax\{J_t(x_t, w_t)\}$ is non-empty and ascending is the content of Topkis [1998][Theorem 2.8.1, pp 78]. Finally, if $W(x)$ is compact, then from Topkis [1998][Theorem 2.8.3(a), pp 78], $W_t^*(x)$ is a sub-lattice of $\mathbb{R}_+$, with a least and a greatest element, both of which are increasing in x. □

[10] Supermodular functions have very attractive properties, in that they are closed under summation and pointwise limits. In fact, these functions form a convex cone which is closed under a variety of useful operations. These properties can be further extended to stochastic versions of supermodular functions. See Athey [2002] for a full development of these properties in a static setting, and Smith and McCardle [2002] for its usage in a dynamic programming framework.

Proposition 6.4.3. *Under Assumption 6.3.8 and 6.3.9, $V_t(x)$ is decreasing in t, for every $x \in X$*

Proof. We provide a proof by induction. $V_T = \max_{w_t \in W(x)} (\Pi(x, w)) \geq 0 = V_{T+1}$. Let the hypothesis hold for some $k < T$ i.e. $V_k(x) \geq V_{k+1}(x)$. Consider now

$$\begin{aligned} V_{k-1} &= \Pi(x, w) + \delta \mathbb{E}[V_k(x - w + R)] \\ &\geq \Pi(x, w) + \delta \mathbb{E}[V_{k+1}(x - w + R)] \\ &= V_k \end{aligned}$$

□

6.4.2 The effect of Risk aversion on optimal decisions

We turn now to understanding the implications of moving away from a risk neutral setting (implicitly assumed by means of profit maximizing) to a setting where risk does play a role. To be precise, we now consider an economic agent (farmer) whose objective is to maximize expected utility from profit i.e. $U(\Pi(x, w))$ where U is a strictly increasing and concave function. There has been little work, apart from the analysis in Knapp and Olson [1996] (see references therein for related work), on groundwater management in a non-risk-neutral setting and the implications of risk aversion (and separately, inter-temporal substitution preferences) for management of groundwater. Knapp and Olson [1996] consider the problem in a recursive utility framework, and show that the optimal policies vary significantly when compared to a risk neutral setting (and show that inter-temporal elasticity of substitution plays a far larger role than does risk aversion).

We consider instead a simpler setting, with a conventional utility function, with the drawback of an inability to distinguish between risk aversion and inter-temporal substitu-

tion. On the other hand, as we indicate below, we have not come across any literature, including Knapp and Olson [1996], which deals with an analysis of which of the properties proved in Theorem 6.4.2 hold in the setting just considered. For instance, while the work in Knapp and Olson [1996] is, in part, a continuation of Knapp and Olson [1995], there is no clear analysis of the changes induced in the structural properties of the models considered. For instance, it is not clear (as we show below) that many of the properties survive this simple (strictly concave) transformation.

While it is evident that the maximizer cannot change by means of this alteration in the objective function, new difficulties arise as a result of this operation. A major cause for the difficulties encountered is that supermodularity, unlike concavity, is a *cardinal* property i.e. it is not preserved under monotonic transformations. In others words, $g(f)$, where g is strictly increasing and f is supermodular, is guaranteed to be supermodular only if g is also convex. Further, an ordinal generalization of supermodularity, quasi-supermodularity (developed in Milgrom and Shannon [1994]) does not lead to any operational benefits in most cases, since other than strictly increasing transformations of supermodular functions, it is difficult to verify quasi-supermodularity of a function in practice, since there is no differential characterization of this property (see the discussion in [Topkis, 1998, pp 60-61]).

In particular, while quasi-supermodularity is sufficient for preservation of monotonicity (in common with supermodularity), it is unsuited in particular for dynamic analyses since it is not closed under a variety of operations, summation and point-wise limits being two of the more prominent ones for our application. Thus, even if $U(\Pi)$ is quasi-supermodular, it is not true that the resulting J_t would also be, since the sum of a supermodular and a quasi-supermodular function need not be supermodular.

It turns out however that when restricted to the log function, some pleasing properties are retained. In particular, the following is true:

Claim 6.4.4. $\Pi(x, w)$ is log super-modular (log s.p.m)

Proof. A function f is log sp.m if $\log f$ is s.p.m i.e. if $\dfrac{\partial^2 f}{\partial u \partial v} \geq 0$ [Topkis, 1998, pp 64]which implies $f f_{uv} \geq f_u f_v$. In our case, the verification exercise involves $\Pi\Pi_{xw} \geq \Pi_x \Pi_w$. $\Pi_x = cw$, $\Pi_w = B^{'} - c(\bar{x} - x)$, $\Pi_{xw} = c$. Thus, the condition to be verified yields $c\left(B - B^{'}w\right) \geq 0$, which holds only if $B \geq B^{'}w$. That this condition holds for (i) $B(w) = aw - bw^2$ (ii) aw (iii) $ln(w + D)$, $D = 1$[11] and (iv) $B(w) = 1 - \exp(-aw)$[12] is easily seen. These five benefit functions more than span the range of the empirical functions used (the most popular of which is that in (i)). □

Replacing $\Pi(x, w)$ in (6.8b) with $U(\Pi(x, w))$, noting that $\dfrac{\partial U(x,w)}{\partial x} = \dfrac{\Pi_x}{\Pi} \geq 0$ (whenever Π is bounded away from 0), it is evident that Theorem 6.4.2 is applicable, which is the content of

Proposition 6.4.5. *Theorem 6.4.2 is applicable if Assumption 6.3.5 is replaced with Claim 6.4.4, and Assumption 6.3.8 with Assumption 6.3.10.*

Remark 6.4.6. (Finiteness of J_t) We comment now on the assumption of boundedness and non-negativity of Π. Note first that, for the purposes of Theorem 6.4.2, in particular, for the maximization operation, it is required that either Π be finite or that $\mathcal{C} = \{(x, w); w \in W(x)\}$ is compact and $J_t(x, w)$ is upper semi-continuous (both of which of course yield finiteness of J_t and the existence of a maximizer), which latter we assume in Theorem 6.4.2. For these purposes, it is *not sufficient* that Π is non-negative, an approach which is the most common when using more conventional approaches to dynamic programming for unbounded rewards.

Remark 6.4.7. (Boundedness of Π) For the particular case of the logarithmic utility function, since the log function is unbounded both above and below, it is not sufficient that X is

[11] For $B > B^{'}w$ for this last functional form, it is required that $\log(1 + w) \geq \dfrac{w}{1+w}$. However, that this holds is evident from the logarithmic inequality $\dfrac{w}{1+w} \leq \log(1 + w) \leq w$.

[12] In this case, $B - wB^{'} = 1 - \exp(-aw)\,(1 + aw)$. That this is non-negative is evident from the following inequality: $e^{-x}(1 + x) < 1$, with $x = aw$.

compact, since $\Pi = 0$ is always a possibility. However, this can be easily dealt with in a very general manner, as follows: let Π be uniformly bounded below (when Π is finite or defined over a compact set, this is trivial but otherwise, it is not) by $B > -\infty$. Consider now the function $\tilde{\Pi} = \Pi + (1 + B)$. It is evident that $\tilde{\Pi} > 0$ and further, that replacing Π with $\tilde{\Pi}$ yields the same optimal decision. Therefore, there is no loss of generality in assuming Π to be uniformly bounded away from 0 in both verification of log supermodularity as well as in Claim 6.4.4.

Remark 6.4.8. (Log supermodularity) Knapp and Olson [1996] consider an identical problem to that in Knapp and Olson [1995], using a recursive utility framework. Their main assumption regarding the function Π[13] ([Knapp and Olson, 1996, A.4., pp 1007]) is

$$\sigma \Pi \Pi_{xw} \geq \Pi_x \Pi_w \tag{6.11}$$

It is obvious that, for $\sigma \geq 1$, this condition is implied by the log supermodularity of Π. On the other hand, they use a value of $\sigma < 1$, for which this condition in fact is more demanding. Using the integral of the linear demand function used in Knapp and Olson [1995] i.e. using $B(w) = aw - bw^2$, and the standard cost function $C(x, w) = c(x)w = c(\bar{x} - x)w$, for the condition in (6.11) to hold, it is necessary that $a \geq \left(\frac{\sigma - 2}{\sigma - 1}\right) bw$. That this condition is stronger than that implied by $B_w(w) \geq 0 \implies a \geq 2bw$ for any $\sigma < 1$ is evident. They appear not to recognize the condition in (6.11) as a (stronger) form of log supermodularity.

6.5 Cost function accounting for local cones of depression

We begin first with the risk neutral setting, and prove that the use of a seemingly more complicated cost function does not complicate the analysis. Indeed, there is no change in

[13] Which is somewhat different from ours due to the differential nature of timing of uncertainty as well as the addition of a second state variable, surface water supplies.

the structural results obtained above.

Recall from (6.7) that the cost function is

$$C(x, w) = c\,(\bar{x} - x + aw)\,w \tag{6.12}$$

while the objective function is

$$\Pi(x_t, w_t) = B(w_t) - c\,(\bar{x} - x + w)\,w \tag{6.13}$$

From $\dfrac{\partial^2 \Pi}{\partial x \partial w} = c > 0$, it is evident that Π is supermodular, and so too that $\Pi_x = cw \geq 0$. This leads immediately to

Proposition 6.5.1. *For the net benefit function in* (6.13)*, Theorem 6.4.2 is directly applicable.*

Further, if we assume finiteness of Π, then it is straightforward that Proposition 6.4.3 is directly applicable, which is

Proposition 6.5.2. *For the net benefit function in* (6.13)*, Proposition 6.4.3 is directly applicable.*

We turn next to characterizing the properties of the decision problem when one uses a (strictly) concave transformation of net benefits. The issues confronted with this seemingly negligible change are substantial, as indicated above. However, even with this cost function, we may

Claim 6.5.3. Π in (6.13) is log supermodular, under Assumption 6.3.10.

Proof. Consider $\Pi_w = B^{'} - c\,(\bar{x} - x + 2w)$, $\Pi_x = cw$, $\Pi_{wx} = c$, which yields

$$\begin{aligned}\Pi\Pi_{xw} - \Pi_x\Pi_w &= c\left(B - wB^{'}\right) - c^2 w\,(-w) \\ &= c\left(B - wB^{'}\right) + c^2 w^2 \\ &\geq 0 \quad \text{iff } B > wB^{'}\end{aligned}$$

which is valid for the variety of functions for B exhibited in the proof of Claim 6.4.4. □

As before, this leads to the following

Proposition 6.5.4. *Theorem 6.4.2 is applicable if Assumption 6.3.5 is replaced with Claim 6.5.3, and Assumption 6.3.8 with Assumption 6.3.10.*

6.6 Accounting for impact of pumping on marginal cost

We show in Section 6.6.1 that all of Properties (a)-(d) hold for model 6.6. This is in contrast to model 6.5 and 6.7, for which we are able to show that only (b) holds. In a related model, Knapp and Olson [1995] are able to show that only (c) holds – and that with the help of an additional condition on c . The underlying *assumption* of Knapp and Olson [1995] is the supermodularity of $B(x-y)-C(x,x-y)$ in (x,y), where $B(w)$ is the benefit of withdrawing w units of water, but this condition sometimes fails for model 6.5; however, for 6.6, this supermodularity condition always holds, thereby enabling Property (c) to hold.

A common way to prove results such as (a)-(d) is via a dynamic programming analysis that simultaneously proves a result about the concavity of the value function. We use this approach, with the slight variation that it is the concavity of the value function *plus* an expression involving the current state that is used in the analysis.

With this cost function, structural properties such as the concavity of the optimization problem in each period have not been established.

EXAMPLE. Under model 6.9, suppose that $B(w) = 5w - w^2/4$, $c(x) = 10 - x$ and $T = 2$. Then it is an easy calculation that $w_2^*(x_2) = 2\max[0, x_2 - 5]$. So, for $5 \leq x_2 \leq 10$, we have $x_2 - w_2^*(x_2) = 10 - x_2$. Thus, Property (c) does not hold.

Continuing analysis of this example, it follows that $V_2(x_2) = (x_2 - 5)^2$ if $5 \leq x_2 \leq 10$, and $V_2(x_2) = 0$ if $x_2 \leq 5$. Suppose that $\delta = 1$ and $R_2 = 0$ with probability 1. Then, for example, $V_1(6) = \arg\max_{w_1} H(w_1)$, where

$$H(w_1) = \begin{cases} B(w_1) - c(6)w_1 + (6 - w_1 - 5)^2 & \text{if } w_1 \in [0, 1), \\ B(w_1) - c(6)w_1 & \text{if } w_1 \in [1, 5]. \end{cases}$$

It is easy to check that the function $H(w_1)$ is not concave in the region $w_1 \in [0, 5]$ (in fact, it has two maximizers $\{0, 2\}$). Thus, Property (a) does not hold.

6.6.1 Analysis

Recall from 6.6 our cost function $C(x, w) = \int_{z=x-w}^{x} \gamma(z)dz$. From the definitions of V_t and G, we can write $V_t(x_t) = \max_{w_t \geq 0} U_t(x_t, w_t)$, where

$$U_t(x_t, w_t) \quad = \quad B(w_t) - \int_{z=x_t-w_t}^{x_t} \gamma(z)dz + \delta E[V_{t+1}(x_t - w_t + R_t)] \ . \tag{6.14}$$

For any x_t, let $w_t^*(x_t) = \arg\max_{w_t \geq 0} U_t(x_t, w_t)$.

In Theorem 6.6.1 below, part (ii) shows that the problem facing the decision maker in each period is the maximization of a concave function, which is Property (a). Part (iii) establishes two properties of the optimal decision in each period – that the optimal withdrawal quantity increases in the groundwater stock, and that the groundwater stock in the next period is increasing in the groundwater stock in the current period – which are Properties (b) and (c), respectively. We prove these results by showing that a modification of V_t exhibits the concavity property – which is the content of part (iv).

Theorem 6.6.1. *(i) $V_t(x_t)$ is increasing in x_t for each $t \in \{1, \ldots, T+1\}$,*

(ii) $U_t(x_t, w_t)$ is concave in w_t for any x_t for each $t \in \{1, \ldots, T\}$,

(iii) w_t^* satisfies $w_t^*(x_t) \leq w_t^*(x_t + \epsilon) \leq w_t^*(x_t) + \epsilon$ for any x_t and $\epsilon > 0$ for each $t \in \{1, \ldots, T\}$,[14] *and*

(iv) $V_t(x_t) + \int_{z=0}^{x_t} \gamma(z)dz$ is concave in x_t for each $t \in \{1, \ldots, T+1\}$.

Proof. We proceed by backward induction. As a base case, note that V_{T+1} is a zero function and trivially satisfies (i) and (iv). We assume, as an induction hypothesis, that (i) and (iv) hold for V_{t+1}, where $t \in \{1, \ldots, T\}$, and prove (i)-(iv) for t.

The inductive step for (i) is immediate from (6.14), i.e.,

$$U_t(x_t, w_t) \quad = \quad B(w_t) - [\Gamma(x_t) - \Gamma(x_t - w_t)] + \delta E[V_{t+1}(x_t - w_t + R_t)] \ ,$$

where $\Gamma(y) = \int_{z=0}^{y} \gamma(z)dz$ is concave, and V_{t+1} is assumed increasing.

To make an inductive step for (ii), we write (6.14) as

$$\begin{aligned} & U_t(x_t, w_t) \\ = \ & [B(w_t) - \Gamma(x_t)] && (6.15) \\ & + (1-\delta)\Gamma(x_t - w_t) && (6.16) \\ & - \delta E[\Gamma(x_t - w_t + R_t) - \Gamma(x_t - w_t)] && (6.17) \\ & + \delta E[V_{t+1}(x_t - w_t + R_t) + \Gamma(x_t - w_t + R_t)] \ . && (6.18) \end{aligned}$$

We now explain why each of the lines (6.15)-(6.18) is individually a concave function of w_t. This is certainly true of (6.15), as B is assumed to be concave. As regards (6.16), we recall

[14] This implies that the derivate of $w_t^*(x_t)$ with respect to x_t should belong to $[0, 1]$ if it exists. Also, this derivative exists almost everywhere since w_t is bounded, increasing and continuous.

the assumption that $\gamma(z)$ is decreasing in z. It follows that $\Gamma(y) = \int_{z=0}^{y} \gamma(z)dz$ is concave in z, and so $\Gamma(x_t - w_t)$ is concave. As regards (6.17), the assumption that $\gamma(z)$ is convex implies that $\partial[\Gamma(v + R_t) - \Gamma(v)]/\partial v = \gamma(v + R_t) - \gamma(v)$ is increasing in v. This implies that $[\Gamma(v + R_t) - \Gamma(v)]$ is convex in v, as is $E[\Gamma(x_t - w_t + R_t) - \Gamma(x_t - w_t)]$. Finally, it is the inductive hypothesis for (iv) that ensures that (6.18) is concave in w_t. We have now established the inductive step for (ii).

We now turn to the inductive step for (iii). This follows from Topkis [1998](pp 78, Thm 2.8.1), which applies to our problem since we are seeking to maximize a function of the form $f_t(x_t, w_t) = U_t(x_t, w_t) + \Gamma(x_t) = B(w_t) + g_t(x_t - w_t)$, where both B and g_t are concave functions, taking g_t to be the sum of lines (6.16)-(6.18). Recall that a function h is said to be supermodular if $h(a \vee b) + h(a \wedge b) \geq h(a) + h(b)$ (where $a \vee b$ and $a \wedge b$ denote maximum and minimum). So, f_t is a supermodular function of (x_t, w_t) because, for any positive ϵ and δ,

$$\begin{aligned} & f_t(x_t + \epsilon, w_t + \delta) + f(x_t, w_t) \;\geq\; f_t(x_t + \epsilon, w_t) + f(x_t, w_t + \delta) \\ \iff \quad & g_t(x_t + \epsilon - w_t - \delta) + g_t(x_t - w_t) \;\geq\; g_t(x_t + \epsilon - w_t) + g_t(x_t - w_t - \delta) \end{aligned}$$

and this is implied by the concavity of g_t. Thus, Topkis's theorem tells us that $w_t^*(x_t + \epsilon) \geq w_t^*(x_t)$. To show $w_t^*(x_t + \epsilon) \leq w_t^*(x_t) + \epsilon$, which is equivalent to $(x_t + \epsilon) - w_t^*(x_t + \epsilon) \geq x_t - w_t^*(x_t)$, we need only make the same application of Topkis's theorem, but take the function to be maximized $\varphi_t(x_t, y_t) = B(x_t - y_t) + g_t(y_t)$, making the identification of $y_t = x_t - w_t$.

Finally, to make the inductive step for (iv), we consider the equation involving (6.15)-(6.18), move $\Gamma(x_t)$ to the left hand side, and replace w_t with $w_t^*(x_t)$. This gives

$$V_t(x_t) + \Gamma(x_t) \quad = \quad B(w_t^*(x_t)) + g_t(x_t - w_t^*(x_t)) \; .$$

Note that since $w_t^*(x_t)$ maximizes $U_t(x_t, w_t) = B(w_t) - \Gamma(x_t) + g_t(x_t - w_t)$ with respect to w_t, it is also the value of w maximizing

$$B(w) + g_t(x_t - w) \tag{6.19}$$

subject to $w \geq 0$. Since B and g_t are concave, the expression in (6.19) is jointly concave in x_t and w. Thus,

$$B(w_t^*(x_t)) + g_t(x_t - w_t^*(x_t)) \quad = \quad \max_{w \geq 0} \quad B(w) + g_t(x_t - w)$$

is concave in x_t (See, for example, Section 3.2.5 of Boyd and Vandenberghe [2004]). Therefore, we conclude that $V_t(x_t) + \Gamma(x_t)$ is concave, proving (iv). □

We note that an alternative way to prove Property (c) is to use a sample path argument: if $x^1 < x^2$ and $x^1 - w^1 > x^2 - w^2$ hold for some period t, then replacing w^1 with $w^1 + x^2 - x^1$ and w^2 with $w^2 - x^2 + x^1$ results in a solution that is not worse than the original solution – this occurs because of the supermodularity of $B(x - y) - C(x, x - y)$ in (x, y).

The following result shows that with more periods to go, the optimal decision is more conservative in the extraction of water (Property (d)), and implies in particular that the optimal extraction quantity is bounded above by the myopic withdrawal quantity (which corresponds to the last period).

Corollary 6.6.2. *$V_t(x)$ is submodular in (t, x). Furthermore, for any x and $t \leq T$, $w_t^*(x) \leq x_{t+1}(x)$.*

Proof. We remind the reader

$$U_t(x_t, w_t) \quad = \quad B(w_t) - [\Gamma(x_t) - \Gamma(x_t - w_t)] + \delta E[V_{t+1}(x_t - w_t + R_t)] \ ,$$

where $\Gamma(y) = \int_{z=0}^{y} \gamma(z)dz$, and recall

$$V_t(x_t) \quad = \quad \max_{w_t \geq 0} \; U_t(x_t, w_t) \; ,$$

where $V_{T+1}(x_{T+1}) = 0$. To show the submodularity of $V_t(x)$ in (t, x), we need to show, for any $1 < t \leq T+1$ and any pair of x^1 and x^2 satisfying $x^2 < x^1$,

$$\Delta_t(x^1, x^2) \quad = \quad V_t(x^1) + V_{t-1}(x^2) - V_{t-1}(x^1) - V_t(x^2) \;\leq\; 0.$$

As a base case, note that $\Delta_{T+1}(x^1, x^2) \leq 0$ holds since V_{T+1} is a zero function, and V_T is an increasing function (Theorem 6.6.1(a)). We proceed with a backward induction on t, by assuming $\Delta_{t+1}(x^1, x^2) \leq 0$ for any $x^2 < x^1$, where $1 < t \leq T$. Let $w_t^1 = w_t^*(x^1)$ and $w_{t-1}^2 = w_{t-1}^*(x^2)$.

Case $x^1 - w_t^1 \geq x^2 - w_{t-1}^2$. From the definitions of w_t^1 and w_{t-1}^2, we obtain

$$V_{t-1}(x^1) = \max_{w \geq 0} U_{t-1}(x^1, w) \geq U_{t-1}(x^1, w_t^1) \quad \text{and} \quad V_t(x^2) = \max_{w \geq 0} U_t(x^2, w) \geq U_t(x^2, w_{t-1}^2) \; .$$

Thus,

$$\begin{aligned} \Delta_t(x^1, x^2) \quad &\leq \quad U_t(x^1, w_t^1) + U_{t-1}(x^2, w_{t-1}^2) - U_{t-1}(x^1, w_t^1) - U_t(x^2, w_{t-1}^2) \\ &= \quad \delta E \Big[V_{t+1}(x^1 - w_t^1 + R_t) + V_t(x^2 - w_{t-1}^2 + R_{t-1}) \\ &\qquad -V_t(x^1 - w_t^1 + R_{t-1}) - V_{t+1}(x^2 - w_{t-1}^2 + R_t) \Big] \; . \end{aligned}$$

Since R_t and R_{t-1} have an identical distribution, we apply the induction hypothesis to conclude the rightmost expression above is at most zero. It follows that $\Delta_t(x^1, x^2) \leq 0$.

Case $x^1 - w_t^1 < x^2 - w_{t-1}^2$. Recall $x^1 > x^2$. It can be shown easily that both $w_{t-1}^2 + x^1 - x^2$ and $w_t^1 - x^1 + x^2$ are nonnegative. By using an argument similar to the one used

in the previous case,

$$V_{t-1}(x^1) \geq U_{t-1}(x^1, w_{t-1}^2 + x^1 - x^2) \quad \text{and} \quad V_t(x^2) \geq U_t(x^2, w_t^1 - x^1 + x^2) \ .$$

Then, from the definition of U_t,

$$\begin{aligned} \Delta_t(x^1, x^2) &\leq U_t(x^1, w_t^1) + U_{t-1}(x^2, w_{t-1}^2) - U_{t-1}(x^1, w_{t-1}^2 + x^1 - x^2) - U_t(x^2, w_t^1 - x^1 + x^2) \\ &= B(w_t^1) + B(w_{t-1}^2) - B(w_{t-1}^2 + x^1 - x^2) - B(w_t^1 - x^1 + x^2) \ , \end{aligned}$$

where the right-hand most expression is at most zero by the concavity of B and $w_{t-1}^2 \leq \min\{w_{t-1}^2 + x^1 - x^2, w_t^1 - x^1 + x^2\} \leq \max\{w_{t-1}^2 + x^1 - x^2, w_t^1 - x^1 + x^2\} \leq w_t^1$.

Thus, we complete the induction step, implying the submodularity of V_t for each t. Now, since $V_t(x)$ is submodular in (t, x), by applying the Topkis's theorem used in the proof of Theorem 6.6.1, we obtain that $w_t^*(x)$ is nonincreasing in t. □

Remark 6.6.3. In the paper, we have assumed that $C(x, w) = \int_{z=x-w}^{x} \gamma(z)dz = \Gamma(x) - \Gamma(x - w)$.

It is evident (from an inspection of the proof of Theorem 6.6.1) that, for our analysis, the differentiability of Γ is not required. Indeed, it is sufficient if $C(x, w) = \Gamma(x) - \Gamma(x - w)$, with Γ increasing and concave and $\Gamma(v + R) - \Gamma(v)$ convex in v for any $R > 0$.

We also note that the proof of Theorem 6.6.1 can be generalized in a straightforward manner to the case where the $\{R_t\}$ are no longer independent or identically distributed (i.i.d). However, the i.i.d property is required for the proof of Corollary 6.6.2.

6.6.2 Connection to the Fighter Ammunition Problem

We comment on how the results of this paper is related to the problem of how an aircraft should spend its missiles over a finite time horizon to maximize the number of enemy planes

shot down, known as the *fighter problem* (see Bartroff and Samuel-Cahn [2011]; Bartroff *et al.* [2010]; Weber [1985]). For the so-called "invincible fighter" case, Bartroff *et al.* [2010] and Bartroff and Samuel-Cahn [2011] have established, for the continuous-time Poisson arrivals and either discrete or continuous ammunition, some of the desirable properties analogous to Properties (b), (c) and (d).[15] It turns out that a special case of our model is the discrete-time[16] continuous-ammunition version of the fighter problem with the invincible fighter, and thus our results in Properties (b), (c) and (d) continue to also hold in this context, providing an alternate proof of the results shown in Derman *et al.* [1975].

Remark 6.6.4. (Log supermodularity) We remark that the profit function associated with this problem, (6.13), is not log supermodular, as a result of which, there is no result analogous to Proposition 6.5.4.

6.7 Effects of the Time horizon

We have already proved the following properties regarding the effect of time horizon. For the cost functions in (6.5) and (6.7), we prove only the (obvious) result that V_t are decreasing in t (and finite, for each t, since each J_t is finite under the assumption of either finiteness of Π or of upper semi-continuity of J_t for each t and the compactness of $\mathcal{C}$) in Propositions 6.4.3 and 6.5.2. For the cost function in (6.6), however, we are able to prove, in addition, convergence of the optimal policy function $w_t(x)$.

A natural next step therefore is to ask the following questions regarding the limit functions. To fix matters, consider the following two equations, which may or may not be well

[15] Properties (b), (c) and (d) in this paper correspond to properties (B), (C) and (A) of Bartroff *et al.* [2010] and Bartroff and Samuel-Cahn [2011], respectively.

[16] The special case that directly follows from Section 6.6.1 assumes that exactly one enemy plane appears in each period. If an enemy plane instead appears with a certain probability, then the proofs of the paper can be easily modified to show that the same set of properties hold.

defined, at this stage (these are the infinite horizon analogues of (6.8a) and (6.8b)):

$$V(x) = \max\{J(x,w); x \in X, w \in W(x)\} \tag{6.20a}$$

$$J(x,w) = \Pi(x,w) + \delta\mathbb{E}[V(x-w+R)] \tag{6.20b}$$

1. Does V_t converge?
   (a) If so, does it converge to the Bellman equation, (6.20a)?
2. Does the optimal policy function $w_t^*(x)$ converge?
   (a) If so, does the limit function inherit monotonicity?
   (b) Finally, does the limit function maximize the right hand side of (6.20b)?

We answer each of them in turn. We begin with a series of brief remarks on convergence of the value function and the one-period return function as a prelude to answering the questions posed above.

Remark 6.7.1. We have proved that V_t, for all three cost functions, is decreasing in t, increasing x, and is finite (in fact, uniformly bounded). It is therefore immediate (from monotone convergence of functions) that $\exists V$ s.t. $V_t \downarrow V$ and further, that the limit function V is increasing in x.

Remark 6.7.2. It is evident from (6.8b) that if $V_{t+1} \leq V_t$ then so too is J_t i.e. $J_{t+1} \leq J_t$. This may be proved by a straightforward induction on t. Further, either due to finiteness of Π or Assumption 6.4.1 and compactness of $\mathcal{C}$, it is evident that J_t is uniformly bounded over $\mathcal{C}$. Therefore, $\{J_t\}$ is a monotone decreasing and (uniformly) bounded sequence. It therefore follows that $\exists J$ s.t. $J_t \downarrow J$ for each $(x,w) \in \mathcal{C}$.

Remark 6.7.3. We have already shown, in Claims 6.4.4 and 6.5.3, that the profit functions in (6.10) and (6.13) are log supermodular. Remarks 6.7.1, 6.7.2 are therefore directly applicable to these formulations.

We make two final assumptions before we embark on our major result for this section.

Assumption 6.7.4. *$J_t(x, .)$ is continuous in w on $W(x)$.*

Assumption 6.7.5. *$\mathcal{C} = \{(x, w); x \in X, w \in W(x)\}$ is a sub-lattice of $\mathbb{R}^2_+$.*

We now state our main

Theorem 6.7.6. *[Heyman and Sobel, 2003, Thm 8-16, pp 427] Under Assumptions 6.3.5, 6.3.10, 6.3.12, 6.3.13, 6.3.14, 6.3.15, 6.7.4, 6.7.5 and Remarks 6.7.1, 6.7.2, and 6.7.3, $V(x)$ (defined in (6.20a)) is non-decreasing in x and $\exists w(x)$ non-decreasing in x which satisfies $V(x) = J(x, w(x))$ for $x \in X$.*

Proof. We provide an outline of the proof, split into two steps. The first step (Theorem 8-14 (pp 418-419)) involves proving that (a) $V_t \downarrow V$ and that (b) V satisfies equation (6.20a). (a) is Remark 6.7.1. A proof of (b) begins by noting that since each V_t is bounded and increasing in x, by the monotone convergence theorem for integrals, we have that

$$\lim_{t\to\infty} \mathbb{E}\{V_t[\tilde{X}(x, w)]\} = \mathbb{E}\left\{V\left[\tilde{X}(x, w)\right]\right\}. \tag{6.21}$$

From this, it is immediate that

$$V_t(x) = \max\{J_t(x, w)\} \geq \max\{J(x, w)\}$$

which implies

$$\lim_{t\to\infty} V_t(x) = V(x) \geq \max\{J(x, w)\}$$

The proof will be complete if the inequality is proved the other way i.e. if it is established that $V(x) \leq \max\{J(x,w)\}$. To prove this, consider

$$V(x) \leq V_t(x) = \sup\{J_t(x,w); w \in W(x)\}$$

and taking limits, observe that

$$V(x) \leq \lim_{t\to\infty} [\sup\{J_t(x,w); w \in W(x)\}]$$

The proof is complete if

$$\lim_{t\to\infty} [\sup\{J_t(x,w); w \in W(x)\}] = \sup\left[\lim_{t\to\infty} \{J_t(x,w); w \in W(x)\}\right]$$

i.e. if the operation lim and sup can be interchanged. From Assumptions 6.7.4 and 6.3.14, it follows that $J_t(x,.)$ converges uniformly in w, for each x, to $J(x,.)$ and finally, equation (6.21) justifies interchange of integral and limit. Thus, the interchange of lim and sup is valid.

The next step, to prove that $W^*(x)$ is ascending in x and the greatest element of $W^*(x)$, $w^*(x)$, is increasing in x, is a direct consequence of Theorem 6.4.2, and is plainly a result of the supermodularity of J. □

Remark 6.7.7. Theorem 6.7.6 addresses question 1. We remark that nothing yet has been said about convergence of the optimal policy i.e. regarding $w_t^*(x) \to w^*(x)$ where w^* is presumably increasing in x and is the maximizer of the right hand side of equation (6.20b). We address in turn questions 2, 2a and 2b. To begin addressing question 2a, we observe that if the limit function w^* exists, it must be increasing in x. The question thus to be addressed is 2. In the case of the cost function in equation (6.6), we have already proved that w_t^* is decreasing in t. Thus, using the upper bound $\bar{X}$ for x, we have that $\{w_t^*\}$ is a

decreasing, uniformly bounded sequence, which must converge. For the remaining two cost functions, we have been unable to prove that w_t^* is decreasing in t, which is a sufficient condition for convergence. Thus, convergence is not assured for the remaining forms of the cost function, *including the conventional cost function.*

Finally, for the cost function in equation (6.6), we turn to see if question 2b can be answered in the affirmative. The main issue is whether the limit function is identical with the maximizer (or belongs to $W^*(x)$) identified in Theorem 6.7.6. The only sufficient condition to ensure that this is so, Heyman and Sobel [2003](Thm 8-15, pp 425-426), relies on the concavity of the value function, and so is not applicable in our case. We have therefore been unable to find a sufficient condition (or provide a direct proof) that the limiting function (to which the optimal policy function converges), in the case of cost function in (6.6), solves the infinite horizon problem in equation (6.20b).

6.8 Stationary Distribution of Stock

We turn now to understanding the conditions under which the Markov Chains generated by the dynamic decision problem for each of the three cost functions (in (6.5), (6.6) and (6.7)) converge to an invariant distribution. Almost all of the literature on renewable resource management, including Knapp and Olson [1995, 1996], relies strongly on the monotonicity of the "reinvestment function" (see below for details) i.e. of the transition equation in terms of the reinvestment function in the current stock, and uses functional analytic methods of stochastic stability, epitomized by Hopenhayn and Prescott [1992]; Stokey and Lucas [1989] in the Economics literature for characterization of the invariant distribution. This approach has at least two major drawbacks: it almost always is based on weak convergence of probability measures and therefore, typically, requires assumptions of compactness of the state space, leading invariably to bounded shocks (an undesirable artifact, in many applications) and (b) is very difficult to generalize to non-monotonic systems.

There is an alternative approach, popularized in Economics in Stachurski [2009], which uses a combination of the function analytic and probabilistic approaches, in particular drawing on the fundamental work of Meyn and Tweedie [1993]. A major advantage is that non-monotonic systems may be handled by this method with almost equal facility. We explore this avenue in particular for the optimal policies for the cost functions in equations (6.5) and (6.7). Since the use of this approach in Economics is relatively recent, and since this approach has not been used (to our knowledge) in Resource Economics, we provide a more detailed outline of the method of verifying the conditions sufficient for the theorem to be applicable. We illustrate here the use of a very powerful theorem, applicable to *both monotonic and non-monotoninc* Markov Chains, under a set of mild assumptions which are likely satisfied in a variety of natural resources. A major advantage of this approach is in allowing the researcher to look beyond monotonic systems, which in many cases in Resource Economics are an artifact of model assumptions rather than any underlying feature of the natural (or economic) system being studied[17].

6.8.1 The Setup

The generic transition equation, which (following Stachurski [2009]) we label the "Stochastic Recursive System" (S.R.S, henceforth), is

$$X_{t+1} = f(x_t) + R_{t+1} = F(x_t, R_{t+1}) \tag{6.22}$$

where for this section, we will assume $R_t \sim^{iid} \Phi$, where Φ is continuous and assigns strictly positive probability to every subset of $\mathbb{R}_+$. We continue with some notation, used

[17] There is yet a third approach, applicable to Monotone Markov Chains, developed in detail in Bhattacharya and Majumdar [2007] based on a "splitting condition" on the Markov Chain, developed in Bhattacharya and Majumdar [1999]. The major advantage of this approach is that irreducibility is not required of the Chain. See also Bhattacharya and Majumdar [2010] for an illustration of limit theorems for Markov Chains in such a setting. We remark that the method of proof we use *yields* irreducibility of the Markov Chain, as a result of which such a generalization is not needed for the current application.

below to elucidate ideas. Denote by $\mathcal{P}(X)$ the set of all (Borel-) probability measures on the state space X, a (Borel-) subset of $\mathbb{R}_+$ and denote by ψ any invariant distribution of the S.R.S in equation (6.22). P is the *stochastic kernel*[18] for the S.R.S in (6.22), defined for $B \in \mathcal{B}(X)$ (set of Borel subsets of X) and $x \in X$, as

$$P(x, B) = \int \mathbb{I}_B \left[F(x, z)\right] \Phi(dz)$$

with $\mathbb{I}_B(.)$ a indicator function for the set B[19]. Given the stochastic kernel P, an associated linear operator, the Markov Operator $M : \mathcal{P}(X) \to \mathcal{P}(X)$ maybe defined as

$$\Phi M(B) := \int P(x, B)\Phi(dx).$$

The Markov operator is therefore a linear operator mapping distribution functions from $\mathcal{P}(X)$ to $\mathcal{P}(X)$. Let $X_0 \sim \Psi$, where $\Psi \in \mathcal{P}(X)$ and denote the distribution of $(X_t)_{t\geq 0}$ as Ψ_t. Then it is the case that the recursion $\Psi_{t+1} = \Psi_t M$ holds, which yields, by an inductive argument, the infinite-dimensional version of the usual Markov Chain identity: $\Psi_{t+1} = \Psi M^t$.

The operator M also acts on *functions*, and is used to define the expectation of any function w.r.t the stochastic kernel i.e. $Mh(x) := \int h(y)P(x, dy)$ $x \in X$. A final point regarding the operator M is that it acts on distributions (alternatively measures) to the *left* and functions to the *right*[20]. We now collect a few formal definitions of the properties

[18] A stochastic kernel is also known as Markov kernel or transition probability function, analogous to the finite state space case, where it is called the transition probability matrix.

[19] To understand the definition of the stochastic kernel, observe that $P(x_t, B) = \mathbb{P}\left(F(x_t, R_{t+1}) \in B\right) = \mathbb{E}\mathbb{I}_B \left[F(x_t, R_{t+1})\right]$. Further, two important properties of this kernel are: for every $x \in X$, $P(x, .)$ is a probability measure and for $B \in \mathcal{B}(X)$, $P(B, .)$ is a measurable function on X. Note finally that the definition of a stochastic kernel above is adapted to the setting of a S.R.S. The general definition of a stochastic kernel is as a family of probability measures. Notationally, a general stochastic kernel is represented as $P(x, dy) \in \mathcal{P}(X)$, $(x \in X)$.

[20] A more precise, and less confusing, usage is to view the Markov operator, M as acting on functions and its *adjoint*, M^*, defined on the space of probability (alternately, finite signed) measures on X, acting on measures (distributions). This is the approach used in Bhattacharya and Majumdar [2007], in the Economics

of P and M, indicating also sufficient conditions for these properties to be fulfilled. We list these properties of M and P since they are integral in the proofs of both our main theorems. More details on these properties, including those regarding the recursions stated above and the linearity of M, may be found in many standard texts on Markov processes (cf. Stachurski [2009](Chapters 8, 9.2 and 11), Stokey and Lucas [1989](Chapter 12), Meyn and Tweedie [1993], Feller [1971] among many others).

Definition 6.8.1. (Feller Property) A Markov Operator M is said to possess the "Feller Property" if it maps bounded, continuous functions into bounded continuous functions.

Remark 6.8.2. (Sufficient condition for Feller property) It is straightforward to prove [Stachurski, 2009, Lemma 11.2.3, pp 258] that if the map $F(x, R)$ is continuous in x on X for each R, then a Markov operator satisfies the Feller property. In fact, for F linear (affine) in x, it is sufficient that F is bounded, since continuity is immediate in this setting.

Definition 6.8.3. (Iterates of M) We have already indicated that $Mh(x)$ may be interpreted as the expectation of the function h under the kernel P. We now illustrate the connection between the t^{th} iterate of M applied to h and the same iterate applied to a distribution Ψ. Note first that since $P^t(x, dy)$ may be interpreted as the distribution of X_t given $X_0 = x$, it is immediate that $M^t h(x)$, defined as

$$M^t h(x) := \int h(y) P^t(x, dy) = \mathbb{E}\left(h(X_t) | X_0 = x\right) \ (x \in X)$$

may be interpreted as a conditional expectation. Then, it can be shown (Stachurski [2009] §9.2 (Thm 9.2.15)) that

(a) $\Phi(Mh) = (\Phi M)(h) = \int \left[\int h(y) P(x, dy)\right] \Phi(dx)$

literature. However, we do not follow such function analytic preciseness since the usage we indicate is more common in the probability literature.

(b) By induction, $\Phi(M^t h) = (\Phi M^t)(h) = \int \left[\int h(y) P^t(x, dy)\right] \Phi(dx)$

Thus, $\Phi(M^t h)$ is merely an (unconditional) expectation.

Definition 6.8.4. (Drift to Small Set) The kernel P satisfies *drift to a small set*[21] if $\exists v \geq 1$, $v : X \to \mathbb{R}_+$, $\alpha \in [0, 1)$ and $\beta \in \mathbb{R}_+$s.t

$$Mv(x) \leq \alpha v(x) + \beta$$

and all sub-level sets of v are small.

Definition 6.8.5. (Aperiodic Kernel) A kernel P is said to be *aperiodic* if it has a $(\nu, \epsilon) -$ *small* set C with $\nu(C) > 0$.

Definition 6.8.6. (Irreducible kernel) A kernel P is said to be $\mu -$ *irreducible* with $\mu \in \mathcal{P}(X)$ if $\forall x \in X$, $B \in \mathcal{B}(X)$ s.t. $\mu(B) > 0$, $\exists t \in \mathbb{N}$ s.t. $P^t(x, B) > 0$. If P is irreducible an arbitrary $\mu \in \mathcal{P}(X)$, then it is called *irreducible*.

Definitions 6.8.5 and 6.8.6 are simply infinite dimensional analogues of the classical definitions for finite dimensional Markov Chains.

Definition 6.8.7. (Global Stability) Viewing $(\mathcal{P}(X), M)$ as a dynamical system, *global stability* corresponds to the existence of a unique fixed point of the dynamical system[22].

Definition 6.8.8. (Stability) Let Ψ^*be an invariant distribution. Denoting by $ib(X)$ the set of increasing and bounded functions on X, *stability* of Ψ^* is taken to mean

$$\forall \Psi \in \mathcal{P}(X),\ h \in ib(X) \quad \left(\Psi M^t\right)(h) \to \Psi^*(h) \text{ as } t \to \infty \tag{6.23}$$

[21] Let $\nu \in \mathcal{P}(X)$, $\epsilon > 0$. A set $C \subset B(X)$ is called $(\nu, \epsilon) -$ *small* for P if $\forall x \in C$, it is the case that for $A \in \mathcal{B}(X)$, $P(x, A) \geq \epsilon \nu(A)$. If this condition holds for *some* ν and $\epsilon > 0$, then the set C is called *small*.

[22] The most common metrics used on $\mathcal{P}(X)$ are either the total variation or the Fortet-Mourier metrics.

This cryptic condition may be easily understood as simply stating that the convergence above is a stronger form of convergence than weak convergence of the (measure) distribution Ψ. To see this, note that we have already defined $(\Psi M^t)(h)$ as an expectation, in Definition 6.8.3. Therefore, weak convergence of Ψ would imply the convergence in (6.23) for $h \in ibc(X)$. However, noting that $ibc(X) \subset ib(X)$, it is immediate that the convergence in (6.23), which holds for $h \in ib(X)$, is stronger than weak convergence.

6.8.2 Main Results

In the case of our (or any) dynamic system, the transition equation maybe written

$$X_{t+1} = x_t - w_t + R_t$$

Consider now a (stationary) policy $w(x_t)$, possibly sub-optimal and non-monotonic in x_t. Under this policy, the Markov Chain that results maybe written as

$$X_{t+1} = x_t - w(x_t) + R_t = G(x_t) + R_t = F(x_t, R_t) \tag{6.24}$$

The fundamental questions concern (a) existence of atleast one invariant distribution to the Markov chain generated by the policy $w(x_t)$ and (b) uniqueness of the invariant distribution.

In the case of the cost functions in (6.5) and (6.7), the optimal policy $w(x_t)$ is increasing in x_t but it is *not* the case that $G(x) = x - w(x)$ is also increasing. Thus, the S.R.S $F(x_t, R_t)$ in equation 6.24 is not increasing in x on X and results on stability of monotone Markov Chains are not applicable. We indicate, following Stachurski [2009]§11.3.5, a constructive method of proof which relies on far simpler assumptions than those for monotone Markov Chains, dispensing in particular with the assumption of compactness of X, in particular, that $X = \{x; a \leq x \leq b,\ a, b \in \mathbb{R}_+\}$.

We first state the major theorem of this section, and then indicate how the assumptions required are satisfied in the case of the S.R.S in 6.24.

Theorem 6.8.9. *[Stachurski, 2009, Thm 11.3.36, pp 292] If the Stochastic Kernel P is aperiodic, irreducible and satisfies drift to a small set, then the system $(\mathcal{P}(X), M)$ is globablly stable with a unique stationary distribution $\Psi^* \in \mathcal{P}(X)$.*

This theorem maybe found proved in [Meyn and Tweedie, 1993, Thm 16.0.1].

The verification of the conditions of 6.8.9 crucially depends on the following

Assumption 6.8.10. *The function $G(x)$ in equation 6.24 is continuous, satisfies*

$$G(x) \leq \alpha x + c \tag{6.25}$$

with $\alpha \in [0, 1)$, $c \in \mathbb{R}_+$, $R_t \sim \Phi$, with ϕ a density which is strictly positive on $\mathbb{R}_+$ and $\mathbb{E}(R_1) < \infty$.

We now indicate why this assumption is reasonable in the context of our set up. First, note that $0 < w(x_t) \leq x_t$ is necessary for assumption 6.8.10 to be satisfied for $c = 0$ (with which we work, since it is not critical that c be positive). Thus, it is *required* that $w(x_t)$ is bounded away from 0 for *all positive values* of x, which of course is a reasonable assumption for any reasonably shallow aquifer. In other words, any agent who has incurred the (not insubstantial) fixed costs of accessing the resource (in the case of groundwater, pump and plumbing; in the case of fishery, capital equipment in the form of boats, nets etc) is unlikely to extract 0 quantity. Second, it is evidently *not sufficient* that this assumption is satisfied.

For a differentiable G, in fact, the assumption 6.8.10 is a condition on the derivative of G i.e. $G^{'} \leq \alpha$. We see now that G is required to be a contraction! For the cost function in (6.6), we have already shown that $\dfrac{\partial w(x)}{\partial x} \in [0, 1]$. If we make a further, mild, assumption that $\dfrac{\partial w(x)}{\partial x}$ is *uniformly* bounded (in x) away from zero, then it is evident that we can set $\alpha = 1 - \inf\left\{\dfrac{\partial w(x)}{\partial x};\ x \in X\right\} < 1$.

When $w(x_t)$ is not differentiable, we see that if $1 > K = sup\{x - w(x); x \in X\}$ then we can set $\alpha = K$. Another alternative is to consider $\inf\left\{\frac{w_t(x)}{x}; x > 0\right\} = \kappa$; if $\kappa > 0$, we can set $\alpha = 1 - \kappa < 1$. For the developments below, we assume one of the above is true[23]. For instance, this condition is naturally satisfied whenever, as in our case, x is bounded. Thus, for all the models considered here, we henceforth make assumption 6.8.10.

We now indicate the chain of reasoning verifying all the properties required of P in Theorem 6.8.9. From Stachurski [2009](pp 293), irreducibility of P follows, while it can be shown easily that every compact subset of X is small for P, from which the aperiodicity of P follows ([Stachurski, 2009, pp 292, exercise 11.34 and 11.28]). Finally, finding a function v satisfying Definition 6.8.4 will suffice to prove drift to a small set. It is evident that $v = x$ does do so[24] . Thus, all the properties required of P in Theorem 6.8.9 are satisfied, which leads to

Proposition 6.8.11. *For the groundwater models defined by cost functions, in equations 6.5, 6.6 and 6.7, the stock of groundwater converges to a unique, invariant distribution* Ψ^*.

We stress that this result is very general in that it depends only on two assumptions: (i) that a stationary policy exists and (ii) under a stationary policy, extraction (not necessarily monotonic) is uniformly bounded away from 0. The former is almost always true while the latter is clearly an assumption, and we argue, a very reasonable one.

[23] It is interesting to note from the literature on extinction of natural resources (for instance Olson and Roy [2000]), this assumption is not equivalent to stating that the stock of resources is bounded away from 0 *almost surely* (pp 194). In other words, the assumption above does not require that the stock be strictly positive. Rather, the assumption indicates that even at very low levels of stock, it is always optimal to extract a non-zero quantity of water, and further, that this quantity is bounded below. Such will always be the case if costs are not "too convex". Given that Inada-like conditions to ensure interior solutions cannot be used in this setting, this is an assumption, albeit a reasonable one .

[24] $Mv(x) = \int v(y)P(x,dy)$. The kernel is $P(x,y) = \phi(y - g(x)) > 0$. Using change of variable, $z = y - G(x)$, $v = y$, we have $\int y\phi(y - G(x))dy = \int (G(x) + z)\,\phi(z)dz \leq \alpha x + \beta$, where $\beta = \int z\phi(z)dz < \infty$. It is immediate that all sublevel sets i.e. set of the form $\{x \in X; v(x) \leq K\}\}$ $K \in \mathbb{R}_+$, are compact.

6.9 Conclusions and Extensions

The analysis in the paper had two major objectives, to

i investigate the implications of using more realistic formulations of cost functions for dynamic groundwater management

ii investigate the changes in structural properties obtained when risk aversion of the economic agent is taken into account.

The cost function was generalized in two directions, accounting for localized cones of depression which potentially increase cost of extraction, and taking into account changes in groundwater stock within a season. With the conventional (simplified) cost function, it was shown that only very few structural properties hold, notably monotonicity of extraction (pumping) in groundwater stock. Quite surprisingly, this simple, and intuitive, result has been rigorously proved here for the first time. It was shown, with the former generalization of the cost function, that extraction is increasing in the current stock. With the latter generalization however, it was shown, in addition, that reinvestment (next periods stock) was increasing in current groundwater stock, and further, that extraction (reinvestment) was decreasing (increasing) over time. In other words, it was shown that a longer horizon (very intuitively) leads to slower extraction and more reinvestment. We stress that our results appear to be the first to show these properties for any form of the dynamic groundwater extraction problem in the literature. Further, we also illustrate, with examples, that many of these results do not hold for the conventional, and simplified, cost function.

In addition, we also show that the net benefit, using the conventional cost function as well as the generalization accounting for cones of depression around the well, is log supermodular. As a result, all of the preceding results are directly applicable. Again, apart from Knapp and Olson [1996], who use a recursive utility framework (with different timing

of uncertainty), ours appears to be the only application explicitly proving structural results for objective functions displaying risk aversion[25].

In much of the literature in Resource Economics which use discrete-time dynamic models, decision problems are formulated in such a way as to yield extraction decisions monotonic in stock, primarily by means of smoothness conditions on the primitives, including concavity of the single-period benefit function and convexity of the feasible set. The reasons for doing so appear to be more related to tractability of the resulting models rather than any underlying feature of the model. Abandoning smoothness and using more realistic assumptions on the primitives in our models, we are able to prove most of the properties previously conjectured in the literature for these class of models.

Finally, many models which deal with the existence of stationary distribution for the stock of resource, are cast in such a manner as to yield "re-investment" (or subsequent periods stock, when uncertainty is resolved prior to the decision on extraction being made) which is increasing in current stock. This allows an immediate appeal to standard results in convergence of monotone Markov Chains. However, moving away from models which impose such strong assumptions (for instance, two of our cost function, both the conventional model as well as the generalization accounting for localized cones of depression, do not satisfy this condition), we illustrate the use of a more general method of proof of global stability. This method does not require that the Markov Chain be monotonic; further, only very mild assumptions on the optimal policy are imposed to yield the result.

Three directions for extension of the current work are immediately evident. First, assumptions of finiteness (of the objective function) permeate the use of supermodularity, which is primarily a restriction of lattice-theoretic methods. In particular, for many Resource Economic applications, lower bounds on utility appear contrived and are an unsatis-

[25] Despite discussion in the Agricultural Economics literature (see for instance Krautkraemer *et al.* [1992], Kennedy *et al.* [1994] among others) about the necessity, in dynamic decision problems, of using objective functions which reflect the beliefs of the modeler regarding the behavior of the agents under consideration.

factory manner of dealing with possible unboundedness. In this setting, use of the weighted sup norm approach, explored in detail in Hernández-Lerma and Lasserre [1999] and increasingly used in Economics (see for instance Stachurski [2009](§12.2.1) and references therein) holds some promise.

Second, in the case of groundwater, an important issue is the prevention of groundwater depletion (or restoration of the groundwater system, if already depleted) by means of regulation. In this context, even if the optimal policy is monotone, it need not be a simple function of the stock. In this setting, it is of some interest to explore, even in the admittedly simplified case of a single user, a variety of possibly sub-optimal but simple policies (such as a linear policy, as in Athanassoglou *et al.* [2011]) and to characterize the performance of such policies in a given setting.

Finally, in the groundwater scenario which motivated this work, crop choice determines water use (and groundwater extraction), and exogenous prices determine crop choices. It is, in this setting, of some interest to understand the conditions under which the variance (variability) in prices influence the evolution of groundwater stock. In other words,the question of whether increases in variability of price leads to increase in extraction of stock is of some importance for policy. Exploring this question in the model framework above, we feel, is both feasible and interesting.

Bibliography

R. M. Aggarwal and T. A. Narayan. Does inequality lead to greater efficiency in the use of local commons? The role of strategic investments in capacity. *Journal of Environmental Economics and Management*, 47(1):163–182, 2004.

C.D. Aliprantis and O. Burkinshaw. *Principles of real analysis*. 1998.

S. Athanassoglou, G. Sheriff, and T. Siegfried. Optimal Mechanisms for Heterogeneous Multi-cell Aquifers. 2011.

S. Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187, 2002.

J. Bartroff and E. Samuel-Cahn. The fighter problem: optimal allocation of a discrete commodity. *Advances in Applied Probability*, 43(1):1–10, 2011.

J. Bartroff, L. Glodstein, Y. Rinott, and E. Samuel-Cahn. On optimal allocation of a continuous resource using an iterative approach and total positivity. *Advances in Applied Probability*, 42(3):795–815, 2010.

R. Bhattacharya and M. Majumdar. On a theorem of Dubins and Freedman. *Journal of Theoretical Probability*, 12(4):1067–1087, 1999.

R.N. Bhattacharya and M. Majumdar. *Random dynamical systems: theory and applications*. Cambridge Univ Pr, 2007.

R. Bhattacharya and M. Majumdar. Random iterates of monotone maps. *Review of Economic Design*, 14(1):185–192, 2010.

P. Billingsley. *Probability and measure*. Wiley-India, 2008.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, United Kingdom, 2004.

G. Brown, Jr. and R. Deacon. Economic optimization of a single-cell aquifer. *Water Resources Research*, 8(3):557–564, 1972.

N. Brozovic, D.L. Sunding, and D. Zilberman. Optimal management of groundwater over space and time. *Frontiers in Water Resource Economics*, 29:109–135, 2002.

H.S. Burness and T.C. Brill. The role for policy in common pool groundwater use. *Resource and energy economics*, 23(1):19–40, 2001.

O. R. Burt. Optimal resource use over time with an application to ground water. *Management Science*, 11(1):80–93, 1964.

O. R. Burt. Economic control of groundwater reserves. *Journal of Farm Economics*, 48(3):632–647, 1966.

O. R. Burt. Temporal allocation of groundwater. *Water Resources Research*, 3(1):45–56, 1967.

O. R. Burt. Groundwater storage control under institutional restriction. *Water Resources Research*, 6(6):45–56, 1970.

U. Chakravorty and C. Umetsu. Basinwide water management: a spatial model. *Journal of environmental economics and management*, 45(1):1–23, 2003.

H. Cramer. *Mathematical methods of statistics*. Princeton Univ Pr, 1999.

C. Derman, G. J. Liberman, and S. M. Ross. A stochastic sequential allocation model. *Operations Research*, 23(6):1120–1130, 1975.

W. Feller. *An introduction to probability theory and its applications., Vol. II.* John Wiley & Sons, 1971.

M. Gisser and D. A. Sanchez. Competition versus optimal control in groundwater pumping. *Water Resources Research*, 16(4):638–642, 1980.

P. Hellegers, D. Ziberman, and E. van Ierland. Dynamics of agricultural groundwater extraction. *Ecological Economics*, 37(2):303–311, 2001.

O. Hernández-Lerma and J.B. Lasserre. *Further topics on discrete-time Markov control processes*. Springer, 1999.

D.P. Heyman and M.J. Sobel. *Stochastic Models in Operations Research: Stochastic Optimization*, volume 2. Dover Pubns, 2003.

H.A. Hopenhayn and E.C. Prescott. Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica*, 60(6):1387–1406, 1992.

J.O.S. Kennedy, J.B. Hardaker, and J. Quiggin. Incorporating risk aversion into dynamic programming models: comment. *American Journal of Agricultural Economics*, pages 960–964, 1994.

K. C. Knapp and L. J. Olson. The economics of conjunctive groundwater management with stochastic surface supplies. *Journal of Environmental Economics and Management*, 28(3):340–356, 1995.

K.C. Knapp and L.J. Olson. Dynamic resource management: Intertemporal substitution and risk aversion. *American Journal of Agricultural Economics*, 78(4):1004, 1996.

P. Koundouri. Current issues in the economics of groundwater resource management. *Journal of Economic Surveys*, 18(5):703–740, 2004.

J.A. Krautkraemer, GC Van Kooten, and D.L. Young. Incorporating risk aversion into dynamic programming models. *American Journal of Agricultural Economics*, 74(4):870–878, 1992.

R. Mendelssohn and M. J. Sobel. Capital accumulation and the optimization of renewable resource models. *Journal of Economic Theory*, 23:243–260, 1980.

S.P. Meyn and R.L. Tweedie. Markov chains and stochastic stability. Communications and Control Engineering Series, 1993.

P. Milgrom and C. Shannon. Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994.

T. Mitra and S. Roy. Optimal exploitation of renewable resources under uncertainty and the extinction of species. *Economic Theory*, 28(1):1–23, 2006.

L.J. Olson and S. Roy. Dynamic Efficiency of Conservation of Renewable Resources under Uncertainty* 1. *Journal of Economic Theory*, 95(2):186–214, 2000.

B. Provencher and O. Burt. The externalities associated with the common property exploitation of groundwater. *Journal of Environmental Economics and Management*, 24(2):139–158, 1993.

B. Provencher and O. Burt. A private property rights regime for the commons: The case for groundwater. *American Journal of Agricultural Economics*, 76(4):875, 1994.

C. Roseta-Palma and A. Xepapadeas. Robust control in water management. *Journal of Risk and Uncertainty*, 29(1):21–34, 2004.

H.L. Royden. *Real analysis.(3rd edn. ed.)*. Macmillan, New York, 1988.

S.J. Rubio and B. Casino. Competitive versus efficient extraction of a common property resource: The groundwater case. *Journal of Economic Dynamics and Control*, 25(8):1117–1137, 2001.

S. J. Rubio and J. P. Castro. Long run groundwater reserves under uncertainty. *investigaciones económicas*, 20(1):71–88, 1996.

T.U. Siegfried. Optimal utilization of a non-renewable transboundary groundwater resource-Methodology, case study and policy implications. 2004.

J.E. Smith and K.F. McCardle. Structural properties of stochastic dynamic programs. *Operations Research*, pages 796–809, 2002.

J. Stachurski. *Economic dynamics: theory and computation*. The MIT Press, 2009.

N. Stokey and R. Lucas. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.

D. M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26(2):305–321, 1978.

D.M. Topkis. *Supermodularity and complementarity*. Princeton Univ Pr, 1998.

Y. Tsur and T. Graham-Tomasi. The buffer value of groundwater with stochastic water supplies. *Journal of Environmental Economics and Management*, 21(3):201–224, 1991.

Y. Tsur and A. Zemel. Endangered aquifers: Groundwater management under threats of catastrophic events. *Water Resource Research*, 40, 2004. 06S20, doi:10.1029/2003WR002168.

R. R. Weber. A problem of ammunition rationing. In F. J. Radermacher, G. Ritter, and S. M. Ross, editors, *Conference report: Stochastic Dynamic Optimization and Applications in Scheduling and Related Field*, page 148, held at University of Passau, Facultät für Mathematik und Informatik, 1985.

V. E. Worthington, O. R. Burt, and R. L. Brustkern. Optimal management of a confined groundwater system. *Journal of Environmental Economics and Management*, 12(3):229–245, 1985.

N. Zeitouni and A. Dinar. Mitigating negative water quality and quality externalities by joint mangement of adjacent aquifers. *Environmental and Resource Economics*, 9(1):1–20, 1997.

N. Zeitouni. Optimal extraction from a renewable groundwater aquifer with stochastic recharge. *Water Resources Research*, 40(6):W06S19, 2004.

Appendix 6.A Comment on Knapp and Olson [1995]: Supermodularity Condition

The paper by Knapp and Olson Knapp and Olson [1995] proves property (c) *assuming* the supermodularity $\kappa(x, y)$ defined by $\kappa(x, y) = G(x, x - y)$ where $G(x, w) = B(w) - C(x, w)$. This supermodularity condition is equivalent to

$$0 \quad \leq \quad -B''(x - y) - \frac{\partial}{\partial x}\frac{\partial}{\partial y}C(x, x - y) \tag{6.26}$$

where $w = x - y$.

Under cost function in (6.5), i.e., $C(x, w) = c(x)w$, we have $C(x, x - y) = c(x)(x - y)$. Thus, $\frac{\partial}{\partial x}\frac{\partial}{\partial y}C(x, x - y) = -c'(x)$. Therefore, the supermodularity condition (6.26) is equivalent to $-B''(w) \geq -c'(x)$, i.e., the degree of concavity in B is larger than the rate at which c is decreasing. Therefore, with cost function (6.5), property (c) is satisfied with this *added* condition. (Note: In the example provided in the paper, we have $B''(w) = -1/2$ and $c'(x) = -1$, *not* satisfying this added condition.)

Under cost model (6.6), i.e., $C(x, w) = \int_{z=x-w}^{x} \gamma(z)dz$, we have $C(x, x-y) = \int_{z=y}^{x} \gamma(z)dz$. Then, $\frac{\partial}{\partial x}\frac{\partial}{\partial y}C(x, x - y) = 0$. This implies, along with the concavity of B, that the condition 6.26 is satisfied. This is an alternate proof of property (c) using an intermediate result of Knapp and Olson [1995]. However, Knapp and Olson [1995] do not speak of Properties (a) and (b).

Comment on the Proof of Theorem 1: Part (iv)

In the proof of part (iv), we use the following fact: Suppose $\psi(x_t, w)$ is jointly concave in (x_t, w). Then, $\overline{\psi}(x_t) = \max_{w \geq 0} \psi(x_t, w)$ is concave. This result follows directly from Section 3.2.5 of Boyd and Vandenberghe [2004].

For convenience, we include the proof of this result here. For any pair of x' and x'', let w' and w'' maximize $\psi(x', w)$ and $\psi(x'', w)$, respectively. Let $\overline{x} = (x' + x'')/2$ and $\overline{w} = (w' + w'')/2$. Then,

$$\overline{\psi}(\overline{x}) \;\geq\; \psi(\overline{x}, \overline{w}) \;\geq\; [\psi(x', w') + \psi(x'', w'')]/2 \;=\; [\overline{\psi}(x') + \overline{\psi}(x'')]/2 \;.$$

Thus, we retain this proof.

Appendix 6.B Proof of Theorem 6.6.1 using Derivatives

For each $t \in \{1, \ldots, T\}$,

(i) $V_t(x_t)$ is increasing in x_t,

(ii) $U_t(x_t, w_t)$ is concave in w_t for any x_t,

(iii) w_t^* satisfies $w_t^*(x_t) \leq w_t^*(x_t + \epsilon) \leq w_t^*(x_t) + \epsilon$ for any x_t and $\epsilon > 0$, and

(iv) $V_t''(x_t) \leq -\gamma'(x_t)$ for any x_t.

Proof. We proceed by backward induction. As a base case, note that V_{T+1} is a zero function and trivially satisfies (i) and (iv). We assume, as an induction hypothesis, that (i) and (iv) hold for V_{t+1}, where $t \in \{1, \ldots, T\}$, and prove (i)-(iv) for t.

From the induction hypothesis (i) on V_{t+1}, $Z_{t+1}(x_t - w_t) = E[V_{t+1}(x_t - w_t + R_t)]$ is increasing in x_t. Since γ is decreasing (by assumption), $\int_{z=x_t-w_t}^{x_t} \gamma(z)dz$ is also decreasing in x_t. Thus, for any w_t, $U_t(x_t, w_t)$ in (6.14) is increasing in x_t. It follows now that $V_t(x_t)$ is increasing in x_t, proving (i) for t.

From (6.14),

$$\frac{\partial}{\partial w_t} U_t(x_t, w_t) = B'(w_t) - \gamma(x_t - w_t) - \delta Z'_{t+1}(x_t - w_t) \quad \text{and} \tag{6.27}$$

$$\frac{\partial^2}{\partial w_t^2} U_t(x_t, w_t) = B''(w_t) + \gamma'(x_t - w_t) + \delta Z''_{t+1}(x_t - w_t) \; . \tag{6.28}$$

By the convexity of γ, we have

$$-\gamma'(x_t - w_t) \; \geq \; -E[\gamma'(x_t - w_t + R_t)] \; = \; Z''_{t+1}(x_t - w_t) \; .$$

Therefore, the second and the third term in the right side of (6.28) satisfy

$$\begin{aligned} \gamma'(x_t - w_t) + \delta Z''_{t+1}(x_t - w_t) &\leq \gamma'(x_t - w_t) - \delta\gamma'(x_t - w_t) \\ &= (1 - \delta)\gamma'(x_t - w_t) \\ &\leq 0 \; , \end{aligned} \tag{6.29}$$

where the last inequality follows since $\delta \in (0, 1]$ and γ is decreasing. Thus, by the concavity of B, we conclude that (6.28) is at most 0, i.e., $\frac{\partial^2}{\partial w_t^2} U_t(x_t, w_t) \leq 0$, proving (ii).

The induction hypothesis (iv) on V_{t+1} (that $V''_{t+1}(x_{t+1}) \leq -\gamma'(x_{t+1})$ for any x_{t+1}) implies

$$E[V''_{t+1}(x_t - w_t + R_t)] \; \leq \; -E[\gamma'(x_t - w_t + R_t)] \; . \tag{6.30}$$

Note that the left-hand-side of (6.30) is the same as $Z''_{t+1}(x_t - w_t)$ from the definition of Z_{t+1}. From (**??**), the right-hand-side of (6.30) is bounded above by $-\sup_w B''(w)$, which in turn is bounded above by $-B''(w_t)$. Thus, (6.30) implies

$$Z''_{t+1}(x_t - w_t) \; \leq \; -B''(w_t) \; .$$

We claim that $B''(w_t) + \delta Z''_{t+1}(x_t - w_t) \leq 0$. To see this, if $Z''_{t+1}(x_t - w_t) \leq 0$, then, by the

concavity of B, $B''(w_t) + \delta Z''_{t+1}(x_t - w_t)$ is the sum of two non-positive terms; otherwise, if $Z''_{t+1}(x_t - w_t) > 0$, then we have $\delta Z''_{t+1}(x_t - w_t) \leq Z''_{t+1}(x_t - w_t) \leq -B''(w_t)$. This completes the proof of the claim. Then, since γ is decreasing, we obtain from (6.28) that $\frac{\partial^2}{\partial w_t^2} U_t(x_t, w_t) \leq 0$, proving (ii).

Now, we will prove (iii). Fix x_t. Part (ii) shows that $U_t(x_t, w_t)$ is concave in w_t. Since Part (ii) has now been established we may proceed by assuming that the partial derivative of U_t with respect to w_t is zero at the optimal withdrawal quantity $w_t^*(x_t)$. (We omit here the case of the boundary solution $w_t^*(x_t) = 0$ for brevity.) ; this case can be argued by considering two intervals $[x_t, x_t + \epsilon_1]$ and $[\epsilon_1, \epsilon]$ separately, where ϵ_1 is the smaller of $\min\{\epsilon' \geq 0 \mid U'_t(x_t + \epsilon', 0) \leq 0\}$)

We first consider the case of $w_t^*(x_t) = 0$. In this case, since $U_t(x_t, w_t)$ is concave in w_t, it follows from (6.27) From (6.27),

Since $U_t(x_t, w_t)$ is concave in w_t, by the first order condition and (6.27), $w_t^*(x_t)$ satisfies

$$\begin{aligned} 0 &= \frac{\partial}{\partial w_t} U_t(x_t, w_t)|_{w_t = w_t^*(x_t)} \\ &= B'(w_t^*(x_t)) - \gamma(x_t - w_t^*(x_t)) - \delta Z'_{t+1}(x_t - w_t^*(x_t)) \ . \end{aligned} \tag{6.31}$$

From (6.27), the above identity and the concavity of B, we obtain

$$\begin{aligned} &\frac{\partial}{\partial w_t} U_t(x_t + \epsilon, w_t)|_{w_t = w_t^*(x_t) + \epsilon} \\ &= B'(w_t^*(x_t) + \epsilon) - \gamma((x_t + \epsilon) - (w_t^*(x_t) + \epsilon)) - \delta Z'_{t+1}((x_t + \epsilon) - (w_t^*(x_t) + \epsilon)) \\ &= [B'(w_t^*(x_t) + \epsilon) - B'(w_t^*(x_t))] + B'(w_t^*(x_t)) - \gamma(x_t - w_t^*(x_t)) - \delta Z'_{t+1}(x_t - w_t^*(x_t)) \\ &\leq 0 \ . \end{aligned}$$

It implies, by the concavity of $U_t(x_t + \epsilon, w_t)$ in w_t, the value of w_t maximizing $U_t(x_t + \epsilon, w_t)$ should be bounded above by $w_t^*(x_t) + \epsilon$, i.e., $w_t^*(x_t + \epsilon) \leq w_t^*(x_t) + \epsilon$.

Now, again from (6.27),

$$\begin{aligned}&\frac{\partial}{\partial w_t}U_t(x_t+\epsilon,w_t)|_{w_t=w_t^*(x_t)}\\ &\quad= \quad B'(w_t^*(x_t))-\{\gamma(x_t+\epsilon-w_t^*(x_t))+\delta Z'_{t+1}(x_t+\epsilon-w_t^*(x_t))\}\ .\end{aligned}\tag{6.32}$$

We want to show the nonnegativity of this expression. Since (6.29) implies

$$\frac{\partial}{\partial x_t}\left[\gamma(x_t-w_t)+\delta Z'_{t+1}(x_t-w_t)\right] \quad\leq\quad 0\ ,$$

it follows that $\gamma(x_t-w_t)+\delta Z'_{t+1}(x_t-w_t)$ is an decreasing function of x_t, showing that

$$\gamma(x_t+\epsilon-w_t^*(x_t))+\delta Z'_{t+1}(x_t+\epsilon-w_t^*(x_t)) \ \geq\ \gamma(x_t-w_t^*(x_t))+\delta Z'_{t+1}(x_t-w_t^*(x_t))\ .$$

Thus, we obtain the nonnegativity of (6.32) since

$$\begin{aligned}&\frac{\partial}{\partial w_t}U_t(x_t+\epsilon,w_t)|_{w_t=w_t^*(x_t)}\\ &\quad\geq\quad B'(w_t^*(x_t))-\{\gamma(x_t-w_t^*(x_t))+\delta Z'_{t+1}(x_t-w_t^*(x_t))\}\\ &\quad=\quad \frac{\partial}{\partial w_t}U_t(x_t,w_t)|_{w_t=w_t^*(x_t)}\\ &\quad=\quad 0\ ,\end{aligned}$$

where the last equality follows since $w_t^*(x_t)$ maximizes $U_t(x_t,w_t)$. which implies that (6.32) is nonnegative, i.e., $\frac{\partial}{\partial w_t}U_t(x_t+\epsilon,w_t)|_{w_t=w_t^*(x_t)}\geq 0$.

To do this, we first claim that (6.32) is bounded below by

$$B'(w_t^*(x_t))-\{\gamma(x_t-w_t^*(x_t))+\delta Z'_{t+1}(x_t-w_t^*(x_t))\}\ .\tag{6.33}$$

To prove this claim, observe that, for any y,

$$Z''_{t+1}(y) \;=\; E\left[V''_{t+1}(y+R_{t+1})\right] \;\leq\; -E\left[\gamma'(y+R_{t+1})\right] \;\leq\; -E\left[\gamma'(y)\right] \;=\; -\gamma'(y)\;,$$

where the first equality follows from the definition of Z_{t+1}, the first inequality follows from the induction hypothesis (iv), and the second inequality follows the convexity of γ and the nonnegativity of R_{t+1}. Thus, by the nonnegativity of $-\gamma'$, we obtain $\gamma'(y)+\delta Z''_{t+1}(y) \leq 0$. (If $Z''_{t+1}(y) \leq 0$, then this sum adds two non-positive numbers; if $Z''_{t+1}(y) > 0$, then $\gamma'(y)+\delta Z''_{t+1}(y) \leq \gamma'(y)+Z''_{t+1}(y) \leq 0$.) Then,

$$\begin{aligned} 0 \;&\geq\; \int_{x_t-w_t^*(x_t)}^{x_t+\epsilon-w_t^*(x_t)} \left[\gamma'(y)+\delta Z''_{t+1}(y)\right] dy \\ &=\; \left\{\gamma(x_t+\epsilon-w_t^*(x_t))+\delta Z'_{t+1}(x_t+\epsilon-w_t^*(x_t))\right\} \\ &\quad -\left\{\gamma(x_t-w_t^*(x_t))+\delta Z'_{t+1}(x_t-w_t^*(x_t))\right\}\;. \end{aligned}$$

Note that the right-hand-side of the above expression is the difference between (6.33) and (6.32). Thus, we prove the claim, i.e.,

$$\frac{\partial}{\partial w_t}U_t(x_t+\epsilon,w_t)|_{w_t=w_t^*(x_t)} \;\geq\; B'(w_t^*(x_t))-\gamma(x_t-w_t^*(x_t))-\delta Z'_{t+1}(x_t-w_t^*(x_t))\;.$$

Now, from the first order condition (6.31), we conclude that the right-hand-side of the above inequality is zero, and thus, we obtain $\frac{\partial}{\partial w_t}U_t(x_t+\epsilon,w_t)|_{w_t=w_t^*(x_t)} \geq 0$.

By the concavity of $U_t(x_t+\epsilon,w_t)$ in w_t, we conclude $w_t^*(x_t+\epsilon) \geq w_t^*(x_t)$. We complete the proof of (iii).

Finally, since

$$V_t(x_t) \;=\; U_t(x_t,w_t^*(x_t)) \;=\; B(w_t^*(x_t))-\int_{z=x_t-w_t^*(x_t)}^{x_t}\gamma(z)dz+\delta Z_{t+1}(x_t-w_t^*(x_t))\;,$$

it follows that

$$\begin{aligned} V_t'(x_t) &= B'(w_t^*(x_t))w_t^{*\prime}(x_t) - \gamma(x_t) + \gamma(x_t - w_t^*(x_t))(1 - w_t^{*\prime}(x_t)) \\ &\quad + \delta Z_{t+1}'(x_t - w_t^*(x_t))(1 - w_t^{*\prime}(x_t)) \\ &= B'(w_t^*(x_t)) - \gamma(x_t) \\ &\quad - \left[B'(w_t^*(x_t)) - \gamma(x_t - w_t^*(x_t)) - \delta Z_{t+1}'(x_t - w_t^*(x_t))\right](1 - w_t^{*\prime}(x_t)) \ . \end{aligned}$$

By the first order condition (6.31), the expression inside the square bracket above is zero, and it follows that

$$V_t'(x_t) = B'(w_t^*(x_t)) - \gamma(x_t) \ .$$

Since B is concave and $w_t^*(x_t)$ is increasing in x_t (by (iii)), we obtain

$$V_t''(x_t) = B''(w_t^*(x_t)) \cdot w_t^{*\prime}(x_t) - \gamma'(x_t) \leq -\gamma'(x_t) \ ,$$

proving (iv). □

Appendix 6.C Technical Appendix

The aim of this appendix is to prove, under very reasonable hypotheses, that (1) V_t is twice differentiable (1) The interchange of expectation and differentiation is justified and (3) the policy function is differentiable.

(1) and (2) are used in proof of all four parts of Theorem 1 and are therefore critical. We also note that, for the purposes of our proof, we do not require that the shocks be *i.i.d*; instead, the only restriction is that the shocks be Markovian, to be made precise below.

In our proof's below, we make assumptions regarding (a) existence and (b) continuity

of V_t''. We note however that our assumptions are weaker than those in, for instance, Tsur and Graham-Tomasi [1991] (who outright assume thrice differentiability of V_t) and are somewhat standard in the literature.

6.C.1 The setup

We will first set up some notation and definitions, before attempting the full justification. Let $(\mathcal{R}, \mathcal{F}, \mu)$ be the probability space under consideration, $X = [\underline{x}, \overline{x}]$ be the state space (evidently a compact subset of $\mathbb{R}$), $W \subset X$ be the "action" space. Denote by $\tilde{X}$ the transition equation for the state, with $\tilde{X} : (X \times W \times \mathcal{R}) \to X$, with $\mathcal{R}$ being the space of "shocks"; thus, $\tilde{X} = \tilde{X}(x_t, w_t, R_t)$. The value function therefore maybe written as $V_t : X \to \mathbb{R}$ and note that

$$\begin{aligned}
\frac{\partial V_{t+1}(x_t - w_t + R_t)}{\partial x_t} &= V'_{t+1} \\
\frac{\partial V_{t+1}(x_t - w_t + R_t)}{\partial w_t} &= -V'_{t+1} \\
\frac{\partial^2 V_{t+1}(x_t - w_t + R_t)}{\partial w_t^2} &= V''_{t+1}
\end{aligned}$$

if the said derivatives exist.

We next define the relevant integrals and begin with the simpler case of *i.i.d* shocks and later indicate the modification required to accommodate *non-iid shocks* . Consider first the second component of the RHS of the recursion in equation (1), $E[V_{t+1}(x_t - w_t + R_t)]$ where the expectation is w.r.t the distribution of the random variable R. Begin by observing that, by assumption, the distribution of R is independent of x_t, w_t, and is, in addition, *i.i.d*. Thus, the integral may be written as follows:

$$\mathbb{E}(V_{t+1}(x_t, w_t; R)) = \int_{\mathcal{R}} V_{t+1}(x_t - w_t + R)\mu(dR) \tag{6.34}$$

We now define

$$Z_{t+1}(x_t, w_t) := \mathbb{E}(V_{t+1}(x_t, w_t; R)) = \int_{\mathcal{R}} V_{t+1}(x_t - w_t + R)\mu(dR) \tag{6.35}$$

and observe that (if the relevant derivatives exist and the interchange is justifiable)

$$\frac{\partial Z_{t+1}}{\partial w_t} = \int_{\mathcal{R}} \frac{\partial V_{t+1}(x_t - w_t + R)}{\partial w_t}\mu(dR) = \int_{\mathcal{R}} -V'_{t+1}\mu(dR) = -\mathbb{E}(V'_{t+1}) \tag{6.36a}$$

$$\frac{\partial^2 Z_{t+1}}{\partial w_t^2} = \int_{\mathcal{R}} \frac{\partial^2 V_{t+1}(x_t - w_t + R)}{\partial w_t^2}\mu(dR) = \int_{\mathcal{R}} V''_{t+1}\mu(dR) = \mathbb{E}(V''_{t+1}) \tag{6.36b}$$

We will also assume, as necessary, that X, W and $X \times W$ are either compact or open.

6.C.2 Differentiation

We will now prove, under weak conditions, that V_t has two derivatives.

First Derivative Note that the a.e. differentiability of $V_t(x_t)$ follows from the monotonicity of $V_t(.)$, which is true by the induction hypothesis. In more detail, from Royden [1988][pp 100], we have that V_t is a.e. (w.r.t μ) differentiable and further, that the derivative $V_t^{'}$ is measurable (w.r.t $\mathcal{F}$).

Second derivative We note that there appear to be no general results, for the discrete-time model with non-concave value function, on first or second differentiability of the value function in terms of the primitives of the model(in contrast with the case for the continuous time model). We therefore take the route of appealing to standard theorems for the first derivative which guarantee existence of the second derivative. Most theorems require some form of boundedness, and we collect these below. If $V_t^{'}$ satisfies any of the conditions below, then V_t' is continuously differentiable.

a $V_t^{'}$ is absolutely continuous on X (Royden [1988][pp 109])

b $V_t^{'}$ is monotonic

c $V_t^{'}$ is of bounded variation

We are thus led to either simply assume the existence of V_t'' or to assume that $V_t^{'}$ satisfies one of the above three conditions.

6.C.3 Interchange arguments

We will first define a local definition of integrability. The relevant probability spaces used below have already been defined above, with $U = (X \times W)$.

Definition 6.C.1. If a function $f : (U \times \mathcal{R}) \to \mathbb{R}$ is $\mathcal{F}$-measurable, for each $u \in U$ and continuous in u for each $R \in \mathcal{R}$ it is said to be "caratheodory".

Definition 6.C.2. If f is caratheodory, and $\forall u \in U$, $\exists$ a function g_u in $\mathcal{L}^1(\mathcal{R}, \mathcal{F}, \mu)$ s.t. $\forall y$ in $N(u)$, an open neighbourhood of u, $|f(y, R)| \leq g(R)$ $\forall R \in \mathcal{R}$, then f is said to be Locally uniformly Integrably bounded.

Remark 6.C.3. Note that *f need not be bounded* in u for it to be *$LUIB$*.

Another important property of *$LUIB$* functions are continuity of the integral. Let

$$h(u) := \int_{\mathcal{R}} f(u, R)\mu(dR) \tag{6.37}$$

where f is *$LUIB$*.

Claim 6.C.4. h is *continuous* in u, where h is defined above.

Proof. Follows from Billingsley [2008][Theorem 16.8(i), pp 212] □

We will first state an important theorem which will allow us to claim, under the conditions above, the validity of the interchange. The theorem maybe found proved, for instance, in Aliprantis and Burkinshaw [1998][pp 193-4, Theorem 24.5], Billingsley [2008][Theorem

16.8,pp 222] or Cramer [1999][pp 67-68], and is an application of Lebesgue's dominated convergence theorem. Define $D_i f(u)$ to be the partial derivative of f w.r.t the i^{th} component of the vector u.

Theorem 6.C.5. *Let h be as defined in equation (6.37), $f : (U \times \mathcal{R}) \to \mathbb{R}$ be Caratheorody. If for each i and each R, $D_i f$ exists and is continuous in u $\forall u \in U$ and if, in addition, $D_i f(u, R)$ is LUIB, then h is continuously differentiable and*

$$D_i h(u) = \int_{\mathcal{R}} D_i f(u, R) \mu(dR) \tag{6.38}$$

Proof. The proof follows from Cramer [1999][pp 67-68]. □

Remark 6.C.6. Note that we do **not** require that **either** of f, $D_i f$ be bounded in u for a given R but we do impose that they are integrable in R for each $u \in U$.

Assumption 6.C.7. *Continuity and integrable boundedness of V'_{t+1}*

(i) V'_{t+1} *is* ***LUIB***

(ii) V'_{t+1} *is continuous in x_t (we note that this assumption is redundant when V''_t exists but is reported here only for completeness)*

Assumption 6.C.8. *Continuity and integrable boundedness of V''_{t+1}*

(i) V''_{t+1} *is* ***LUIB***

(ii) V''_{t+1} *is continuous in x_t*

Claim 6.C.9. Identify u with (x_t, w_t), Z_{t+1} with h, V_t with f, note that V'_{t+1} exists μ-a.e.(by Theorem 1 (i)). Then, under Assumption 1,the interchange in equation (6.36a) is valid i.e we have that

$$\frac{\partial Z(u)}{\partial w_t} = -\mathbb{E}(V'_{t+1})$$

Claim 6.C.10. Identifying u with (x_t, w_t), $\dfrac{\partial Z(u)}{\partial w_t}$ with h, V_t' with f, noting that V_{t+1}'' exists μ-a.e.(by assumption) we have that under Assumption 2, the interchange in equation (6.36b) is valid i.e we have that

$$\frac{\partial^2 Z(u)}{\partial w_t^2} = \frac{\partial\left(\dfrac{\partial Z(u)}{\partial w_t}\right)}{\partial w_t} = \frac{\partial(-\mathbb{E}(V_{t+1}'))}{\partial w_t} = \mathbb{E}(V_{t+1}'')$$

In the claims above, we here make the added assumption that V_t'' is continuous, which cannot directly be justified by an appeal to the II of Littlewood's principles (since without assuming that V_t'' is a.e bounded, it is not possible to apply Luzin's theorem on which the II principle is based).

Remark 6.C.11. We note that all of the results carry over to a non-iid *Markovian* setting. In order to minimise notation, we will simply point out the following: if the distribution of R_t is independent of x_t, w_t and R_t is Markovian i.e. $\mathbb{P}(R_{t+1} \in H | \{R_s\}_{s=0}^t) = \mathbb{P}(R_{t+1} \in H | R_t)$ (where $H \subset \mathcal{R}$), then, letting $Q(R, dR')$ denote the transition function, we may write the integral in (6.34) as

$$\mathbb{E}(V_{t+1}(x_t - w_t + R_t)) = \int_{\mathcal{R}} V_{t+1}(x_t - w_t + R) Q(R, dR') \tag{6.39}$$

All of the results on interchange carry over without any modification (given standard assumptions on $Q(.,.)$. Details on this approach for the current model maybe obtained from Stokey and Lucas [1989][Chapter 8]).

chapter 7

Conclusion and Implications

7.1 Extensions

The five chapters in the dissertation have addressed three different issues. We first indicate a few extensions, both ongoing and future work, which constitute a natural progression from the work undertaken here. We finally conclude by pointing out some of the broader implications of the work undertaken here.

7.1.1 Quantifying climate change impact on agriculture

Chapter 2 has been focused on quantifying the impact of climate change on crop yields, for India. While this analysis, arguably, represents an advance in understanding the distribution of the impact as well as on utilizing sub-seasonal features of monsoon variability, there is an important issue which this analysis, and related literature, has been largely unsuccessful in addressing. First, much of the yield increases associated with the "Green Revolution" have been driven by expansion of irrigation facilities. Yet, both the Agriculture and the Economics literatures (Aggarwal and Mall [2002]; Guiteras [2008]; Kumar and Parikh [2001]; Sanghi and Mendelsohn [2008]) have acknowledged but largely not dealt with the issue. In the case of Agricultural modeling literature, the models used inherently are incapable of accommodating such changes, while in the case of the Economics literature, the omission has been due both to the reduced form nature of the models, as well as due to availability of data. This is one of the reasons for exclusion of irrigated areas[1] from studies even in the developed nations (Deschenes and Greenstone [2007]; Schlenker and Roberts [2009]).

Yet, given that a large part of the developing nations are dependent on irrigated agriculture for food production, it is essential to analyze the relationship between irrigation and

[1] Yet another reason is the debate within the Ricardian framework regarding inclusion of irrigation and its potential, and the interpretation of irrigation in this framework. See Cline [2007] for a discussion and Schlenker *et al.* [2005] for an illustration of the importance of accounting for irrigation in a cross-section setting.

yield in more detail. In particular, the question of importance is the extent to which irrigation can act as a buffer in mitigating the impact of temperature increases. Two difficulties loom prominently with accounting for irrigation in studies of climate change impacts on agriculture; first, and foremost, it is not clear how different irrigation systems will evolve in response to altering climate and second, for the developing nations, detailed unit (district/county) level data regarding current irrigation is unavailable. However, for certain states in India, it is possible to reconstruct series of irrigated areas at the district level, and to pair with a repeated-cross section on groundwater levels, to obtain such a relationship.

7.1.2 Precursors and prediction regarding rainfall extremes

Our analysis in Chapters 3-5 dealt with understanding the nature of changes in extreme rainfall events as well as in initiating the study of large-scale precursors to such events. Yet, these have been mere first steps in a more detailed understanding of the precursors of extreme rainfall events. In particular, there are two challenges which still remain. First, the analysis in Chapter 5 and in related literature, on understanding of larger-scale precursors to rainfall extremes, is nascent and an important methodological challenge faced is extracting spatial patterns from gridded (station) data. Second, relating such large-scale spatial (and spatio-temporal) patterns to climate precursors is a challenging task, and much work remains to be done on this aspect.

Initial results on spatial patterns in extreme events indicate that much of the coherence, in both amplitude of reconstructed changes as well as in direction, exists outside, or at the boundaries, of the "core monsoon" region. This result is intriguing since much work on Indian monsoon is focused on this region, both for extremes as well as for more moderate events. If coherent signals exist outside of this region only, then it not only calls into question the intense focus on this particular region but also raises the possibility of improved prediction in the regions outside the "core monsoon", some of which are also quite flood-

prone.

7.2 Implications of our results

Climate Change Mitigation Policy

The impact of climate change on agriculture is a matter of some debate; yet, for sub-tropical nations such as India, a consensus appears to have developed around a significant negative impact. Given the complexity of the issue, there is not a similar consensus regarding the geographical and population-based distribution of its impact. Attempts, such as those in Chapter 2, to quantify such distributional implications on district-level units are of significant importance in understanding and designing policies to mitigate the adverse impacts of climate change. In the context of the debate regarding mitigation strategies, therefore, this type of analysis provides a basis for regional assessment of agricultural strategies, much of which (for the developing world) involves more investment in better crop varieties and irrigation to buffer increased temperatures.

Yet, the analysis above has been incomplete, in that they do not investigate the population impacts of climate change. Since a given change in temperature affects different economic sectors and individuals in these sectors differently, it is not only important to know the aggregate losses suffered; it is equally important to know *who* bears these losses. A recent study by Jacoby *et al.* [2011] provides a compelling picture of the rural landless bearing the largest proportion of the cost, due to the rising prices for cereals. Further, among the implications of such a study is, at first sight, the counter intuitive one, that

> if the objective is protecting the poor from climate change, a focus on helping farmers may be misplaced.... Nevertheless, a better policy in this respect might be to *reduce the share of rural income* derived from climate sensitive employment

([Jacoby *et al.*, 2011, pp 28], emphasis added). While it is definitely possible to take excep-

tion to parts of the statement, the important point is that the *same* estimate of agricultural impacts can lead to very different policy choices, once the distributional implications are taken into account. To the extent that distributional (both in space and populations) implications are substantial, the exercise of quantifying impacts need not be thought of as guiding policy towards the given goal of sectoral intervention; rather, in the case of developing nations, they may also be viewed as meaningfully allowing a *more optimal choice* of the activities undertaken toward mitigation of such impacts.

Bibliography

P.K. Aggarwal and RK Mall. Climate change and rice yields in diverse agro-environments of India. II. Effect of uncertainties in scenarios and crop models on impact assessment. *Climatic Change*, 52(3):331–343, 2002.

W.R. Cline. *Global warming and agriculture: Impact estimates by country*. Peterson Institute, 2007.

O. Deschenes and M. Greenstone. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, 97(1):354–385, 2007.

R. Guiteras. The impact of climate change on indian agriculture. *Unpublished manuscript*, 2008.

H.G. Jacoby, M. Rabassa, and E. Skouas. Distributional implications of climate change in India. *World Bank Policy Research Working Paper*, (5623), 2011.

K. Kavi Kumar and Jyoti Parikh. Indian agriculture and climate sensitivity. *Global Environmental Change*, 1:147–154, 2001.

A. Sanghi and R. Mendelsohn. The impacts of global warming on farmers in Brazil and India. *Global Environmental Change*, 18(4):655–665, 2008.

W. Schlenker and M.J. Roberts. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594, 2009.

W. Schlenker, W.M. Hanemann, and A.C. Fisher. Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. *The American Economic Review*, 95(1):395–406, 2005.