

Agricultural Adaptation to Drought

By

Emily Kay Burchfield

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Environmental Engineering

May, 2017

Nashville, Tennessee

Approved:

Jonathan Gilligan, Ph.D.

Ralf Bennartz, Ph.D.

Janey Camp, Ph.D.

Amanda Carrico, Ph.D.

George Hornberger, Ph.D.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	vi
LIST OF TABLES	vii
LIST OF FIGURES	viii
Chapter	
1. Introduction	1
1.1. Overview	1
2. Agricultural adaptation to drought in the Sri Lankan dry zone.....	4
2.1. Introduction	4
2.2. Sri Lanka	4
2.3. Methods	6
2.4. Results	10
2.4.1. The D1 systems: Negotiation and reallocation.....	13
2.4.2. System B: Control and experience	14
2.4.3. IH and MH: Institutions and culture.....	15
2.4.4. Wahalkada and Padaviya: History and expansion.....	16
2.5. Discussion	17
2.5.1. Infrastructural access	17
2.5.2. Cross-scale interactions	18
2.5.3. Decentralized resource control	18
2.5.4. Radical reallocation	19
2.5.5. Diversification	19
2.5.6. Monitoring agrowell use.....	20
2.5.7. Farmer perception.....	20

2.6. Conclusion.....	20
3. Dynamics of individual and collective agricultural adaptation to water scarcity	22
3.1. Introduction	22
3.2. Fieldwork.....	23
3.3. Model design	25
3.3.1. Entities, state variables, and scales.....	26
3.3.2. Process overview and scheduling.....	27
3.3.3. Design concepts.....	28
3.3.4. Initialization.....	29
3.3.5. Submodels	30
3.4. Results	31
3.5. Discussion	32
4. Forecasting vegetation health at high spatial resolution	34
4.1. Introduction	33
4.2. Experimental design: Location and data type	34
4.2.1. Experimental variable: Location	35
4.2.2. Experimental variable: Data type.....	36
4.3. Methods.....	38
4.3.1. Data processing and matrix construction	39
4.3.2. Autocorrelation-based validation grid.....	40
4.3.3. Model training and selection	41
4.3.4. Model validation.....	42
4.4. Results	43
4.4.1. Model selection	44
4.4.2. Holdout validation.....	44
4.5. Discussion	49

5. Conclusion.....	51
APPENDIX A: Objective functions	52
APPENDIX B: rSTAN code	54
REFERENCES	56

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grant NSF-EAR-1204685. I am grateful to the NSF for the opportunity to work on such an exciting, interdisciplinary project. Thanks to all members of the ADAPT-SL team and to our Sri Lankan collaborators at the National Building Research Organization. Malaka, your patience and friendship made my time in Sri Lanka unforgettable. Thanks to the faculty and staff that support the American Institute for Sri Lankan Studies (AISLS). The Dissertation Enhancement Grant I received from AISLS significantly improved my research.

I am particularly grateful to Amanda Carrico, George Hornberger, and Jonathan Gilligan. Amanda, thank you for your constant moral support, friendship, and advice. You've shown me that women can succeed and thrive in the academy. You have always encouraged me to go further and reach higher. George, your continued encouragement and thoughtful advice have been invaluable. You have taught me how to teach, mentor, and lead with integrity and honesty. Jonathan, it's impossible to communicate my gratitude for your continued mentorship, encouragement, and moral guidance in a few sentences. I am incredibly lucky to have been able to spend the past five years working with you. I am constantly surprised by the breadth and depth of your knowledge and by the level of enthusiasm and energy you bring to all aspects of your work. If I can teach my future students, as you have, to truly love science, then the significant portion of my life I will devote to my work will be successful.

Thanks to the great graduate students I've been able to collaborate with during my time at Vanderbilt. John, I'll miss having you only an office away. You taught me about how exciting and rewarding successful collaborations can be. Arielle and Nick, thanks for helping me navigate the wild world of qualitative research. Elizabeth, thanks for your friendship, support, and of course, the solid hangs. Kate, thanks so much for joining me on this crazy journey through the world of geospatial Bayes (sigh).

Finally, thanks to Madre, Padre, Zus and Sambro for helping me keep it real. Thanks to Martin for helping me make it through and to Jack for reminding me what it's all for.

LIST OF TABLES

Table	Page
1. Description of the datasets used in the predictor sets	38

LIST OF FIGURES

Figure	Page
1. Water management and agroecological zones of Sri Lanka.....	6
2. Principal components analysis results	11
3. Cultivation during the 2014 drought.....	12
4. Interview site locations	13
5. Predictors of OFC cultivation.....	25
6. Variation in <i>bethma</i> as a function of agrowell ownership for different water levels and objective functions	30
7. Variation in income over time for individual farmers, grouped by agrowell ownership for different reservoir levels and objective functions.....	31
8. Land use in the San Joaquin Valley and Sri Lanka	36
9. Methods overview.....	39
10. Model performance for each data type in California and Sri Lanka as measured by the percent reduction in mean squared error below the lagged EVI model for each location.....	43
11. Correlation between Predicted and Actual EVI in California and Sri Lanka.....	46
12. Performance across values of true measured EVI in California agricultural and in Sri Lanka agricultural land	47

13. Correlation between predicted and actual EVI over time periods in California (A.) and Sri Lanka (B.)	48
14. Percent of pixels with missing data over 23 16-day periods of the year	49

CHAPTER 1

INTRODUCTION

1.1. Overview

Drought is a recurring and complex phenomenon that substantially affects both human and natural systems. On average, drought affects more people and causes more economic damage than any other natural disaster (Wilhite & Vanyarkho 2000). Global circulation models suggest that the spatial extent, likelihood, and severity of drought will increase in the future (Dai 2013). These changes, coupled with rapidly increasing population and shifts in consumer demand, will strain global agricultural systems. To ensure future food and water security, it is essential to understand the factors that boost agricultural resilience to drought.

Differences in climate and soil explain only a fraction of crop failures during drought, suggesting that the relationship between drought and agricultural production is confounded by other factors. Drought arises from an interaction between reduced rainfall (meteorological drought), soil moisture stress (agricultural drought), reduced canal flows or reservoir storage (hydrological drought), and restricted water access caused by economic factors or political power (socioeconomic drought) (Heim 2002). Regions with similar infrastructural, institutional, and physical characteristics may manifest markedly different responses to similar drought events (Swain et al. 2014).

This research combines geospatial, qualitative, and survey data to identify the factors that drove agricultural adaptation to drought during a severe drought that hit rural Sri Lanka in 2014. The work presented in this dissertation interrogates multiple disciplinary, stakeholder, and scalar perspectives to increase understanding of the adaptive capacity of water management systems in Sri Lanka. My research objectives are as follows:

Objective 1: Compare adaptive responses to drought across different water management systems.

Objective 2: Model the factors that drive agricultural decision-making and adaptation.

Objective 3: Support future adaptation by developing a decision support system for water managers.

Drought has particularly severe effects on agricultural systems (Lesk et al. 2016). The complex social and ecological processes that interact to generate agricultural responses to drought include management paradigms and governance, cultivation patterns, decision-making processes, information availability and access, infrastructure, and environmental factors (Meinzen-Dick 2007; Ostrom 2009). A system's adaptive capacity, or the ability of a system to prepare for stresses and changes in advance or adjust and respond to the effects caused by the stresses, emerges from complex interactions between these processes at multiple scales and levels (Smit & Wandel 2006; Engle 2011; Gibson et al. 2000). Adaptive systems have high adaptive capacity and exhibit the potential for structural change (Cash et al. 2006), facilitate coordination and deliberation amongst stakeholders (Lebel et al. 2005), foster social learning through critical self-reflection (Pahl-Wostl et al. 2007), and realign decision-making to natural scales (Moss and Newig, 2010). A community's adaptive capacity is a function of both local processes and the larger systems in which these processes are embedded (Cash et al. 2006; Smit & Wandel 2006). In Chapter 2, I combine geospatial and qualitative data to identify the multi-scalar factors driving local agricultural adaptation to drought. Results suggest that though relatively static factors such as infrastructural capacity and physical environment significantly affect agricultural adaptation, dynamic factors such as control of water supply, perceived risk, community cohesion, and local expertise explain significant variation in the adaptive capacity of agricultural systems.

Chapter 3 presents the results of an agent-based model constructed to explore the dynamics of collective and individual agricultural adaptation to water scarcity. We draw on extensive field-work conducted with paddy farmers in rural Sri Lanka to study adaptations to water scarcity, including switching to less water-intensive crops, farming collectively on shared land, and individually turning to groundwater by digging wells. We explore how variability in climate affects agricultural decision-making at the community and individual levels using three types of decision-making, each characterized by an objective function: risk-averse expected utility, regret-adjusted expected utility, and prospect theory loss-aversion. We also assess how the introduction of individualized access to irrigation water with wells affects community-based

drought mitigation practices. Preliminary results suggest that the growth of well-irrigation may produce sudden disruptions to community-based adaptations, but that this depends on the mental models farmers use to think about risk and make decisions under uncertainty.

Chapter 4 addresses concerns raised by Sri Lankan water managers over the paucity of remotely sensed data available to monitor agricultural health. Cloud cover significantly limits remotely sensed data availability on the island for most of the year. This chapter describes a tool constructed with colleagues that produces short-term forecasts of vegetation health at high spatial resolution, using open source software and data that are global in coverage. The tool automates downloading and processing Moderate Resolution Imaging Spectroradiometer (MODIS) datasets, and training gradientboosted machine models on hundreds of millions of observations to predict future values of the Enhanced Vegetation Index. We compared the predictive power of different sets of variables (raw spectral MODIS data and Level-3 MODIS products) in two regions with distinct agroecological systems, climates, and cloud coverage: Sri Lanka and California. Our tool provides considerably greater predictive power on held-out datasets than simpler baseline models.

Chapter 5 synthesizes findings from this dissertation, discusses the broader impacts of this work, and offers ideas for future work.

CHAPTER 2

AGRICULTURAL ADAPTATION TO DROUGHT IN THE SRI LANKAN DRY ZONE

2.1. Introduction

The determinants of agricultural adaptation to drought can be divided into “structural” and “dynamic” variables. Structural variables are those that are slow to change such as storage capacity, management regimes, long standing hierarchies of power, position within a system, or environmental factors such as soil type or slope. Dynamic factors change quickly, and include the number of water users, politics, information access, collective action, and decision processes. The larger and slowly changing structural factors set the conditions within which the smaller, dynamic processes operate. Conversely, many small, dynamic processes may generate changes in structural variables (Gunderson 2001; Giddens 1984). To capture these cross-scale interactions, I have combined remotely sensed and qualitative data to identify the structural and dynamic determinants of agricultural adaptation in surface water irrigation systems in the Sri Lankan dry zone.

This paper focuses on the processes of agricultural adaptation that took place in rural Sri Lanka in response to a severe drought in 2014. The 2014 drought is estimated to have affected the livelihoods of over one million Sri Lankans. 58 percent of the country had completely insufficient water to cultivate during the 2014 dry season (WFP 2014). I analyzed satellite imagery to measure variations in agricultural responses to drought and identify a subset of agricultural communities with similar structural characteristics (i.e. agroecological region, storage capacity, command area, number of farming families, institutional jurisdiction) but different cultivated extents. I conducted key informant interviews in eight of these communities to identify the factors, both structural and dynamic, that influenced variations in cultivated extent during the drought.

2.2. Sri Lanka

Sri Lanka is an island nation off of the southeastern coast of India. The nation experiences two monsoon seasons annually. The northeast monsoon lasts from October to

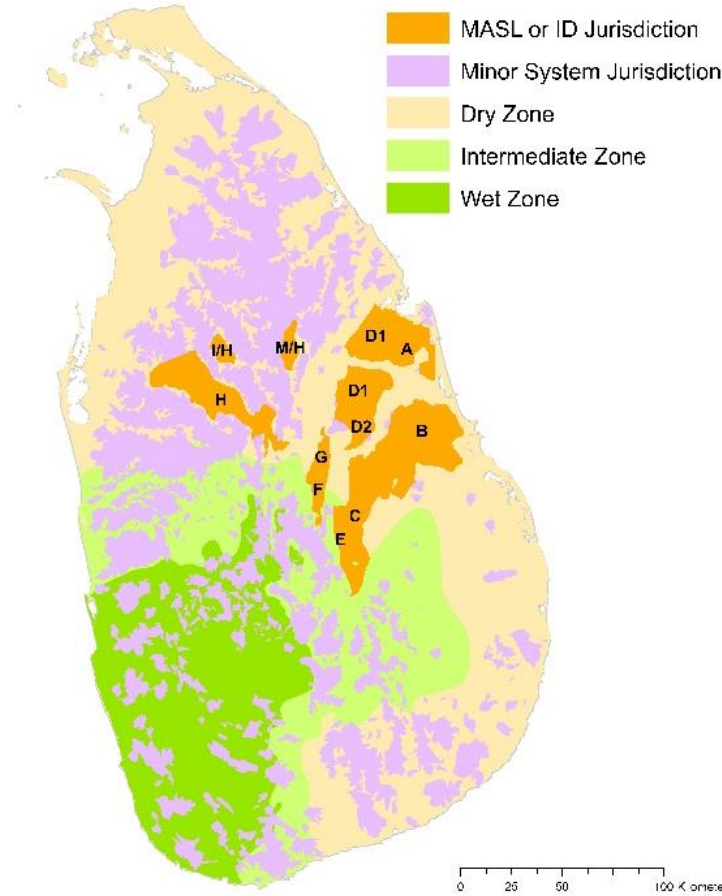
December and brings nearly two-thirds of annual rainfall to Sri Lanka; the southwest monsoon lasts from May to October and brings rain primarily to the southwestern region of the island. This rainfall pattern divides the island into a wet and dry zone (Figure 1) and creates a distinct wet and dry cultivation season.

For over 1,000 years, farmers living in the dry zone have constructed small reservoirs, locally known as *tanks*, to store wet season water for dry season cultivation. Today, the dry zone is dotted with over 11,250 “minor” tank systems (Imbulana et al. 2006). Due to low tank storage capacities, variations in rainfall, and growing population, farmers in these systems frequently experience water scarcity during the dry season (Shah, Samad, Ariyaratne, & Jinapala, 2013). To address these challenges, in the 1960s the Sri Lankan government began construction of a network of massive irrigation systems that diverted the waters of nation’s largest river, the Mahaweli Ganga, through a system of centrally managed reservoirs, hydropower plants, and over 10,000 km of canals (Withananachchi et al. 2014). In the 1970s, the government created the Mahaweli Authority of Sri Lanka (MASL) and charged the institution with the implementation and management of these new “major” irrigation systems (Zubair 2005). The MASL offered perpetual leases to government-owned plots of land in the MASL systems. Farmers who resettled the land received 2.5 acres of paddy land and 0.5 acres of homestead (Takesada et al. 2008). By the end of 2012, the MASL had resettled over 166,000 families onto 250,000 acres of irrigated land (Withananachchi et al. 2014). Today, these irrigation systems contribute significantly to the Sri Lankan economy, producing over 800,000 metric tons of paddy annually (MASL, 2014) and generating enough power to meet 40 % of Sri Lanka’s energy demand (Manthrithilake and Liyanagama, 2012).

Over 40 institutions and legislative acts govern water use in Sri Lanka (Manthrithilake and Liyanagama, 2012). Minor irrigation systems fall under the jurisdiction of the Department of Agrarian Development and are primarily managed by the farmers themselves. The MASL and Irrigation Department (ID) share the management of major irrigation systems. Prior to each season, a group of national officials from the Ceylon Electricity Board, the Department of Agriculture, the ID, and the MASL meet to determine seasonal inflows to each major system reservoir. The group produces a Seasonal Operating Plan (SOP) that specifies the first and last date of water issues for each system, proposed cultivated extents, expected energy generation,

and monthly diversion volumes for each major irrigation system. Within each major irrigation system, water release from reservoirs along main canals is managed by system-level MASL or ID officials. Farmers are grouped by field canal into farmer organizations (10-15 farmers) that are responsible for field-level water rotations and canal maintenance.

Figure 1: Water management regimes and agroecological zones of Sri Lanka.



2.3. Methods

Many studies have used remotely sensed metrics of vegetation health to monitor agricultural responses to drought (Brown et al. 2002; Peters et al. 2002; Thenkabail et al. 2004). I use the Enhanced Vegetation Index (EVI) to measure regional variations in the effects of drought on agricultural vegetation health. The EVI is a strong proxy for rice growth and is highly correlated with both leaf area and vegetation fraction estimates (Gumma, 2011; Huete et

al., 2002; Sakamoto et al., 2005; Small & Milesi, 2013; Xiao et al., 2006). The EVI is measured as:

$$EVI = G \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + C_1 \times \rho_{RED} - C_2 \times \rho_{BLUE} + L}$$

where ρ is atmospherically corrected surface reflectance, L is the canopy background adjustment, and C_1 and C_2 are the coefficients of the aerosol resistance term, which uses the blue band to correct for aerosols in the red band (Huete et al. 2002). EVI values approaching one indicate higher levels of photosynthetic activity.

To first identify double-cropping agricultural communities, I compiled 16-day 250 meter MODIS Terra MOD13Q1.005 EVI imagery from January 2004 to June 2015 into a single spatiotemporal datacube. The EVI time series for each pixel contains information about seasonal changes in vegetation health, land cover, cropping patterns, and a stochastic component. In tropical countries like Sri Lanka, this stochastic component is strongly influenced by cloud cover. Data reduction techniques such as principal component analysis (PCA) can be used to extract phenological information from noisy datasets by separating deterministic processes in lower components and location-specific or stochastic dimensions in higher components (Lasaponara 2006; Small 2012; Eastman 1993). Conceptually, we can think of PCA as rotating a multivariate dataset in such a way that we see the axis along which there is the most variation. This helps us identify patterns in complex datasets, such as a spatiotemporal assemblage of satellite imagery, by reducing the dimensionality of the data with minimal loss of information. The eigenvectors extracted from PCA represent the direction of the axis of the rotation; the eigenvalue is the measure of the variance in that direction. The eigenvector with the highest eigenvalue is the principal component. The first principal components account for much of the variation in the dataset; the last PCs capture directions in which there is very little variation. The components define spatial patterns in the dataset and the temporal coefficients, or loadings, indicate the trace through time of the spatial mode (Anyamba & Eastman 1996). When applied to remotely sensed space-time datacubes, the first PCS typically capture deterministic processes driving land surface processes (i.e. land use, seasonality, water access) and the higher PCs capture stochasticity or processes that vary significantly across space.

Each pixel is associated with a vector \mathbf{x} of EVI values through time. We can represent pixels in a vector space with many dimensions, one for each observation in time. A pixel plots in this coordinate space with coordinates that correspond to its EVI values through time. PCA finds new coordinate systems in this multidimensional space in which the data can be represented without correlation, i.e. the covariance matrix in the new coordinate system is diagonal (Richards 2013). We want to find the linear transformation D^T that transforms the original coordinates \mathbf{x} into a new coordinate system \mathbf{y} such that:

$$\mathbf{y} = D^T \mathbf{x}$$

where each component of \mathbf{y} is a linear combination of the elements of all the elements of \mathbf{x} and the weighting coefficients are the elements of the matrix D^T . We essentially want to find a transformation of the covariance matrix of the original data C_x so that the pixel data in a new \mathbf{y} coordinate space shows no correlation, i.e. the covariance matrix in the new \mathbf{y} coordinate space C_y is diagonal:

$$C_y = D^T C_x D$$

where C_y is a matrix of the eigenvalues of C_x and D is a matrix of the eigenvectors of C_x .

To extract the dominant phenological signals from the noisy dataset, I applied standardized PCA to the unmasked EVI dataset, dropping data from 2014 and 2015 to remove the effects of the drought. The empirical orthogonal functions (EOF) from this analysis represent the data as uncorrelated temporal patterns and the principal components (PCs) represent the spatial distribution of these patterns (Anyamba & Eastman 1996; Eastman 1993).

In our analysis, the third PC captured the contribution of surface water irrigation to variations in vegetation health and showed a strong double-cropping signal through time. I applied various thresholds to the third PC to classify pixels as double-cropped or not and compared this classification to a land use map created by the Sri Lankan Survey Department in 2011. I constructed a receiver operating characteristic (ROC) curve to assess the overall performance of the threshold approach and to determine the appropriate threshold (Hanley & McNeil 1982). The total area under the ROC provides a metric for classification performance. Increasing area indicates increasing performance, with an area of one corresponding to perfect

predictions. Our approach performs well, with a value of 0.80. Using the Youden Index, I found the threshold of the third PC at which the ROC curve is furthest from the line of equity (Fluss et al. 2005). I masked pixels with loadings on the third PC above this value to identify regions in which farmers double-crop, i.e. regularly cultivate their fields during both the wet and dry seasons.

To identify the subset of these double-cropped pixels in which cultivation occurred during the 2014 dry season drought, I used two criteria: total seasonal vegetation production and maximum seasonal EVI. Total seasonal vegetation production is measured as the integral of the smoothed seasonal EVI curve and is a proxy of the amount of biomass produced on a pixel (Jönsson & Eklundh 2004; Rasmussem 1992; Lupo et al. 2007). The maximum seasonal EVI threshold ensures that selected pixels exhibited a greening up during the dry season. Because agricultural fields tend to have peak EVI values great than 0.5, I used this value as the maximum seasonal EVI threshold (Huete et al. 2002; Sakamoto et al. 2005). Missing data caused by cloud contamination were linearly interpolated and smoothed using the Savitzky-Golay filter, a low-pass filter particularly well-suited to noisy data (Savitzky & Golay 1964; Chen et al. 2004). For each double-cropped pixel, I computed the average dry season total vegetation production from 2004 to 2013 and compared it to the 2014 value. Pixels with total seasonal vegetation production greater than one standard deviation below the 10-year pixel average and a maximum seasonal EVI above 0.5 were flagged as those in which farmers were able to cultivate during the drought.

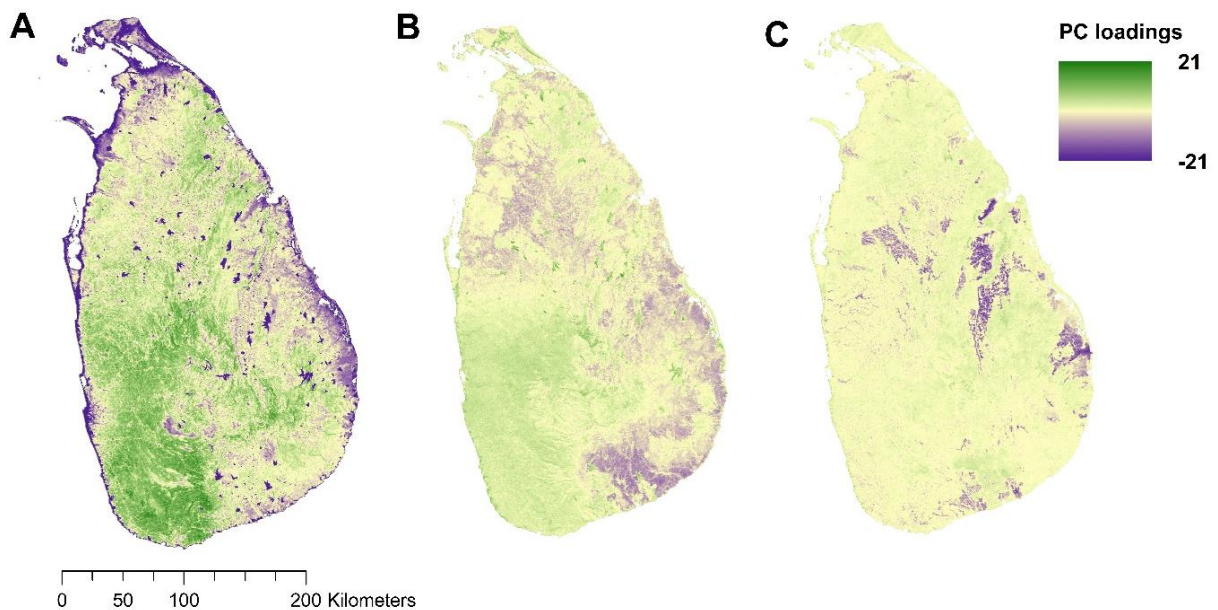
The remotely sensed analysis identified large-scale patterns of agricultural cultivation and served as the foundation for a more detailed analysis of the dynamic factors that affected agricultural adaptation to the 2014 drought. To identify the structural determinants of agricultural adaptation, I linked the results from our remote sensing analysis to a geographic information system (GIS) containing information about the characteristics of agricultural communities, such as agroecological region, storage capacity, command area, number of farming families, institutional jurisdiction, and relative location within the irrigation network. Using this information, I selected four pairs of communities with similar structural characteristics that exhibited different cultivated extents during the 2014 drought. Randomly selected locations in which our larger research project had already established institutional relationships with key

government officials were prioritized in the community selection process. In August 2015, I conducted key informant interviews with local officials, system-level officials, and farmers in each community. Officials included national water managers in Colombo, system-level engineers and water managers, farmer organization officials, and agricultural extension officers. A total of 38 interviews and 4 farmer focus groups were conducted. When interviews could not be conducted in English, they were conducted through a translator. In each interview, I discussed the factors that the interviewee perceived as influencing cultivation during the 2014 drought.

2.4. Results

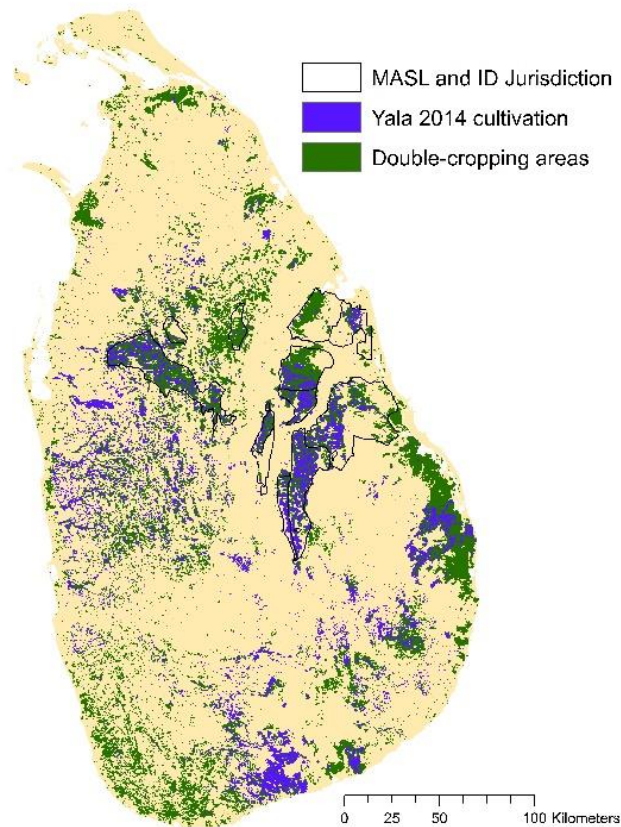
The results of the PCA analysis reveal the spatiotemporal patterns that explain most of the variance in vegetation health in Sri Lanka from 2004 to 2013 (Figure 2). The first PC (41 % of the total variance) captures the contribution of land cover to variations in vegetation health. The second PC (4.4 % of total variance) isolates the seasonal and spatial variations in vegetation health caused by the monsoon, with higher loadings in the wet zone and lower loadings in the dry zone. The third PC (3.1 % of total variance) has very low loadings within the institutional boundaries of the MASL systems and the eigenvector of this PC shows a strong double-cropping signal. This PC captures the contribution of surface water irrigation systems to variations in vegetation health. To identify double-cropped pixels, I applied a threshold to the third PC using the methods described above.

Figure 2: Principal components analysis results: (a) The first PC captures the variations in land cover that explain most of the variance in vegetation health in Sri Lanka. (b) The second PC detects variations in vegetation health attributable to the wet, intermediate, and dry agroecological zones on the island. (c) The third PC shows strong negative loadings within the boundaries of the MASL and ID irrigation systems. This PC captures the contribution of surface water irrigation to the vegetation health variations.



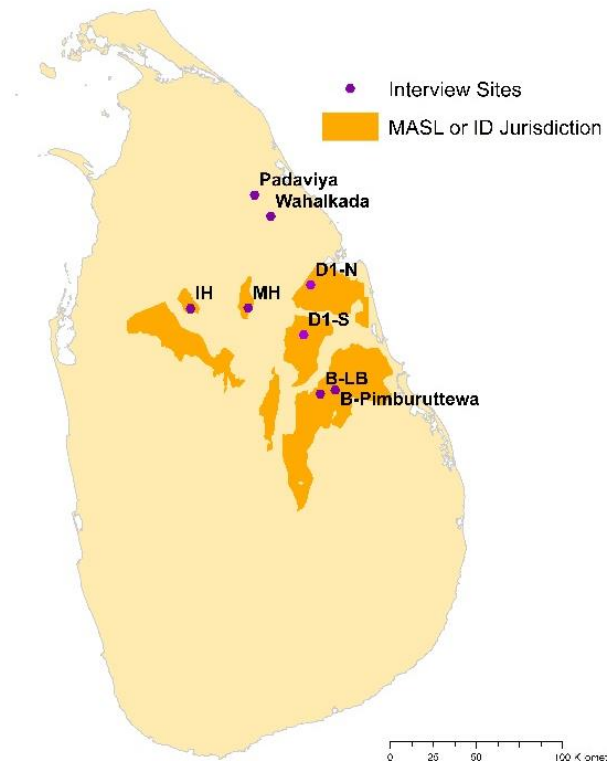
Pixels in which cultivation occurred during the drought (i.e. satisfying the total vegetation production and maximum seasonal EVI criteria) are shown in Figure 3. 45 % of these pixels are located within major system boundaries. Only 25 % of cultivated pixels are located within minor system boundaries, and 65% of these pixels are located in the wet zone. The Survey Department's land use map classified 73% of the identified cultivated pixels as agricultural (slash and burn agriculture known as *chena*, gardens, plantations, or paddy). Of the remaining non-agricultural classified pixels, 16% were classified as roads, forest, or bodies of water located in close proximity to agricultural areas.

Figure 3: Cultivation during the 2014 drought: Green pixels are the regions in which farmers typically double-crop, i.e. cultivate during both the wet and dry seasons. Purple pixels are those in which cultivation occurred during the 2014 dry season drought. Most of these cultivated pixels are located within the southeastern wet zone or are within the jurisdictional boundaries of MASL and ID systems.



To uncover the dynamic, local processes that affected agricultural adaptation, I visited eight dry zone communities (Figure 4) to discuss the 2014 drought with local water managers and farmers. In the following section, I compare these systems to articulate processes described by community members as significantly contributing to agricultural (mal)adaptation during the 2014 drought.

Figure 4: Interview site locations: All sites in which interviews were conducted were located within MASL or ID jurisdiction. Padaviya and Wahalkada fall under the jurisdiction of the ID, but do not receive water from MASL irrigation infrastructure and are considered to be medium-sized rain-fed systems.



2.4.1. The D1 systems: Negotiation and reallocation

At the beginning of the 2014 dry season, the MASL determined that upstream reservoirs were too low to send irrigation water to the D1 systems. Officials warned against cultivation, urging system managers to save limited water in the D1 reservoirs for domestic use. Farmers in both systems staged multiple protests at local ID offices and MASL headquarters in Colombo demanding that officials release irrigation water for paddy cultivation. Farmers argued that they could cultivate paddy and meet domestic water demand if they practiced *bethma*, a traditional drought mitigation technique native to the dry zone. Under *bethma*, permanent field boundaries are temporarily abolished and land is redistributed amongst all farmers who cultivate in the command area. This redistribution process is complex and varies from system to system, but in general, each family receives equal-sized parcels of land regardless of land ownership (de Jong

1989; Thiruchelvam 2010a; Spiertz & de Jong 1992). The total amount of land cultivated by each farmer is temporarily reduced to ensure all farmers in the community have access to limited water supplies.

In the D1 systems, farmers proposed a *bethma* in which head-end farmers would divide their original 2.5 acre fields into half-acre parcels. Each farmer cultivating at the tail-end of the command area would temporarily move to the head-end of the system to cultivate one of the remaining four parcels on each head-end farmer's land. In both D1 systems, this proposed reallocation of land would force tail-end farmers, many of whom belong to the Tamil ethnicity and speak Tamil, to travel over 40 km to cultivate head-end plots which in large part belong to Sinhalese families who speak Sinhalese. Despite these cultural, infrastructural, and physical challenges, farmers still preferred *bethma* to no cultivation.

Local water managers ultimately conceded to farmers' requests to cultivate a small subset of the command area, making it clear that the farmers would bear all risks associated with cultivation. At the end of the season, 19% and 25% of the total command area was cultivated with paddy in Systems D1N and D1S respectively. Farmers attributed this success to increased involvement by local water managers and their own increased water use efficiency. Despite the serious physical and infrastructural constraints faced by D1 farmers, farmers successfully negotiated with officials to cultivate a reduced command area during the drought. Many farmers attributed this success to their political influence as potential voters in the buildup to a national election. After the negotiations were complete, farmers and water managers understood that they alone bore the risk associated with cultivation because the MASL was physically unable to send additional water north. Several farmers and officials said that the high risk increased cooperation in land and water reallocation as well as overall water use efficiency in both systems.

2.4.2. System B: Control and experience

At the beginning of the dry season, System B's main reservoir, Maduru Oya, was filled to half-capacity and the MASL stated that the system would not receive additional inflows for the remainder of the season. To ensure adequate drinking water supplies for this large system, the MASL recommended a 50 % *bethma* in which tail-end farmers would move to the head-end of the system to cultivate. The MASL also advised farmers to grow other field crops such as soy

and maize that are less water intensive than paddy. I visited a community along the left bank of System B in which the cultivated area was reduced during the drought. Farmers in this community agreed to the 50 % *bethma*, though few cultivated the recommended alternative crops, stating that they lacked a local market and necessary agricultural inputs to do so. At the end of the season, these farmers cultivated 59 % of the command area, only 1 % of which was cultivated with other field crops.

I also visited a community in System B in which, according to the remotely sensed results, 100 % of the command area was cultivated during the drought. This community, while technically located in System B, stores irrigation water in a smaller tank downstream of Maduru Oya. Most of these farmers live relatively close to the tank, making it easy for them to monitor their water supply. A group of older farmers inspected the tank's water levels at the beginning of the season and claimed that in the past they had successfully cultivated the entire command area with similar amounts of water. These farmers convinced the other farmers cultivating in the tank's command area to ignore MASL recommendations and cultivate 100 % of the fields with available water. These farmers, like the D1 farmers, took a significant risk and responded by managing water with extreme efficiency. They checked fields daily, monitored water levels, and patrolled for illegal siphons. One farmer proudly stated that by the end of the season the drainage canals were too dry for fish to survive. The experience of a few farmers and the community's control of its water supply facilitated agricultural adaptation to the 2014 drought. Had the farmers listened to MASL recommendations, they would have cultivated only 50 % of their command area.

2.4.3. IH and MH: Institutions and culture

Much of the water delivered to System MH from the wet zone travels through a 73 km feeder canal that transfers water from an upstream reservoir. Local water managers with the ID, the institution responsible for managing water in System MH, claimed that the system's main reservoir rarely received water inflows promised by the MASL because of water poaching along this feeder canal. In response to the structural water scarcity this has caused in System MH, many farmers have installed agrowells and now pump groundwater to irrigate crops. Agrowell irrigation cannot generate sufficient water to cultivate paddy, so many farmers have started cultivating other field crops such as soy, maize, and onions. During the 2014 drought, the

MASL recommended that local water managers not release irrigation water from the main reservoir in System MH to ensure domestic water demands could be met. Because of this restriction, only farmers with access to an agrowell were able to cultivate during the drought. Most of the farmers interviewed had not invested in agrowells and were forced to find employment outside of the agricultural sector.

System IH, a system similar to System MH in terms of command area, storage capacity, and distance from MASL headwaters, showed strong signs of cultivation during the drought. Interviews revealed that farmers in this system received 5,000 acre feet of water from the MASL during the 2014 dry season. The farmers used this water to successfully practice a 50 % *bethma*, 40 % of which included other field crops. Like System MH, System IH receives water from a feeder canal leaving System H. Unlike MH, farmers here do not experience structural water scarcity. When asked to explain the difference in water availability in the two systems, IH officials cited two reasons. The first was institutional fragmentation. Both System MH and IH are managed by the ID, though the MH feeder canal is managed by the MASL while the IH feeder canal is managed by the ID. Officials said that the MASL had little incentive to monitor water overuse along the feeder canal that sent water to a system outside of its jurisdiction. Along the IH canal, however, ID officials actively monitor water poaching and water flow. The second reason cited by officials was the cultural importance of the IH area. System IH also surrounds the city of Anuradhapura, home to some of the most sacred Buddhist sites in Sri Lanka. During the drought “diversions were made ... to address [the] cultural requirement” of the thousands of thirsty pilgrims that temporarily call Anuradhapura home during religious festivals (MASL 2014).

2.4.4. Wahalkada and Padaviya: History and expansion

The remotely sensed analysis revealed radically different cultivated extents in two northeastern minor systems that share similar command areas and storage capacities: Padaviya and Wahalkada. In Wahalkada, farmers surprisingly cultivated 100 % paddy during one of the most severe droughts in recent history. Local farmers attributed their cultivation to the system’s history. Like most of the irrigated communities in the dry zone, farmers were resettled from overpopulated southern cities during the 1960s and 1970s. Today, in most of the dry zone irrigation systems, second and third generation descendants of the original settlers face land

fragmentation, growing population, and increased demand for water (Azmi 2007). Wahalkada's resettlement began relatively late in 1973. At the onset of the civil war in the 1980s, resettlement stopped. After the war ended in 2009, families moved back to the area, but today relatively few families cultivate in the Wahalkada command area. Low water demand allows farmers in the area to cultivate the entire command area even during periods of extreme drought.

Several kilometers down the road in Padaviya, only 19 % of the command area was cultivated during the drought. Padaviya resettlement started in 1954, nearly 20 years earlier than in Wahalkada. Though many farmers left during the war, long-established ties to the region brought them back in the mid-2000s. While Wahalkada's 810 hectare command area supports only 1,185 farming families, Padaviya's 970 hectare acre command area supports over 9,000 families. Overpopulation in Padaviya contributed to water shortages during the 2013 dry season and the 2012 and 2013 wet seasons. These systematic water shortages have pushed many farmers to seek alternative employment. When water managers proposed a 25 % *bethma* during the 2014 drought, many remaining farmers sold their *bethma* plots and abandoned agriculture for the season. The remaining farmers cultivated 19 % of the command area, 100 % of which with crops other than paddy. Despite water managers' efforts to manage water efficiently, at the end of the season water was so scarce that drinking water had to be delivered by truck.

2.5. Discussion

2.5.1. Infrastructural access

The most important driver of cultivation during the 2014 drought was access to MASL irrigation infrastructure. This access facilitated a spatiotemporal transfer of water from the wet season and wet zone to their fields. Despite widespread access to this infrastructure, many MASL farmers questioned whether existing storage capacities were sufficient to support future population growth in the dry zone. The MASL response to these concerns is the construction of the largest reservoir in Sri Lanka, Moragahakanda, which could bring an additional 3500 acres under cultivation (SME Ltd., 2013). Over a thousand families will be displaced to construct this reservoir and thousands more will be resettled into the newly irrigated regions of the dry zone (Ranasinghe 2013). Though infrastructural development is an essential response to changing climate, the expansion of water-intensive agriculture in the dry zone should be executed with

extreme caution. Systems which are located far downstream from MASL headwaters such as the D1 systems already experience severe water scarcity during periods of drought. The overexpansion of agricultural production in the dry zone may push the region past its carrying capacity and gradually erode the adaptive capacity of agrohydrological systems (Holling & Meffe 1996).

2.5.2. Cross-scale interactions

More flexible, democratic, and participatory institutions have been shown to increase adaptive capacity (Engle & Lemos 2010; Gupta et al. 2010; Cash et al. 2006). In most MASL and ID systems, water allocation management is already fairly decentralized. Local water controllers, often farmers themselves, are responsible for opening sluice gates and monitoring water flows at the field-canal level. These water controllers are familiar with canal layouts, canal maintenance needs, and variations in field characteristics (primarily soil type and elevation). This expertise allows them to tailor allocations determined in system offices to local contexts. Farmers organization leaders liaise with water management officials regularly to discuss issues with water access and cultivation. Leveraging this existing organizational structure to increase farmer participation in *system-level* allocation decisions would integrate farmers' unique knowledge of field and canal dynamics into seasonal allocation plans. Similarly, by limiting institutional fragmentation, water scarcity emerging from coordination problems such as those seen in System MH could be avoided in the future.

2.5.3. Decentralized resource control

In System B, local control of water supply allowed farmers to apply their expertise to water release decisions. This autonomy ultimately allowed farmers to achieve 100% cultivation during the drought. Though not always feasible, increasing a community's control of its water supply could be one way of increasing local adaptive capacity. In MASL and ID systems, this may mean creating local tanks to store water as it moves through the system. It would require a reorganization of farmers around these smaller tanks rather than the current organization along field-canal. Though tank-based communities have existed in the dry zone for over a thousand years, this massive restructuring of the MASL infrastructure is not likely. An alternative is to

provide farmers with additional information about water availability to increase their ability to negotiate with system-level and national officials.

2.5.4. Radical reallocation

Bethma is one of the most impressive responses to drought observed in the dry zone. *Bethma* temporarily disrupts the status quo to buffer against inequalities in drought exposure within a community. Despite the prevalence of *bethma*, many farmers doubted that the practice would survive in the future. Land fragmentation has reduced farmers' field size so significantly that many fields can no longer be divided under *bethma*. In addition, the introduction of agrowells has individualized water access, which has encouraged agrowell-owning farmers to opt out of *bethma* and cultivate their entire field using groundwater (Burchfield & Gilligan 2016). At present, system-level officials are mandating that these farmers share their land. As the prevalence of agrowells increases, this mandate is becoming more and more difficult to enforce.

2.5.5. Diversification

Farmers at the majority of the study sites practice paddy monoculture. Though paddy is heavily subsidized, easy to store, and ideal for home consumption, its cultivation is extremely water intensive (Prasanna et al. 2011). At present, farmers have little incentive to cultivate less water intensive field crops such as soy, onions or chilies. There are no subsidy programs and other field crops are much more difficult to store, transport, and sell (Chandrasiri & Bamunuarachchi 2015). The main market for vegetables is located in the center of the island in Dambulla, a significant distance from many dry zone communities. At the end of each season, the Dambulla market is often flooded with a single crop, such as onions or chilies, and farmers are forced to accept extremely low prices. In addition to these market constraints, farmers face infrastructural constraints when cultivating other field crops. In surface water irrigation systems, farmers along the same field canals frequently follow the same water rotation schedule, making it difficult for a single farmer to diverge from the dominant crop planted on that field canal. Increasing support at the national level for agricultural diversification broadens the portfolio of options available to farmers during a drought (Ellis 1998; Lin 2011) and increases an agricultural system's potential to positively respond to a water supply shock (Holling 2001; Liu et al. 2007).

2.5.6. Monitoring agrowell use

In the past, farmers used groundwater predominantly for domestic use. Today, groundwater is increasingly used as a compliment to surface water for irrigation (Villholth and Rajasooriyar 2009). The total number of agrowells in Sri Lanka has increased in the last two decades from zero to more than 50,000 and an estimated 55 percent of farmers in the dry zone now use groundwater to irrigate agricultural fields (Kikuchi et al. 2001). The long-term sustainability of agrowell use is questionable, especially given the fact that in Sri Lanka many of these agrowells are only deep enough to collect surface water drainage (Shah, Roy, Qureshi, & Wang, 2003). The government should carefully monitor agrowell use in the dry zone and study the long-term implications of increased groundwater pumping.

2.5.7. Farmer perception

In systems where farmers bore the risks associated with cultivation beyond command areas proposed by the MASL, farmers engaged in extremely efficient water management practices. Farmers agreed that during normal dry seasons, they rarely monitored fields or water releases because they knew there was sufficient water. During the drought, these farmers applied existing knowledge of efficient water management techniques with rigor. This suggests that though farmers are aware and capable of engaging in efficient water management practices, they lack incentives to manage water efficiently during normal seasons. System-level officials could establish norms and incentives for the farmers to manage water efficiently and to report misuse during normal seasons.

2.6. Conclusion

Despite massive infrastructural and institutional investments in the dry zone over the past 50 years, water scarcity remains a serious problem. Droughts of a serious nature occur every three to four years, while severe droughts occur every ten years (Imbulana et al. 2006). Growing population has increased demand for land and water, causing land fragmentation, landlessness, encroachment, and water scarcity (Azmi 2007). The Sri Lankan population is expected to increase by 15% in the next 30 years, further straining limited water supplies (UN, 2006). Climate scientists predict that farmers will face a decrease in wet season rainfall and an increase in dry season drought in the future (De Silva et al. 2007; Jayawardene et al. 2005; Malmgren et

al. 2003). The demographic, economic, and environmental changes facing Sri Lanka challenge agrohydrological systems around the world. Research exploring how these complex resource management systems respond to water stress is of paramount importance if we are to meet growing demands in an increasingly stressed physical environment.

The findings of this project suggest that though structural factors such as water management regime boundaries, infrastructural capacity, relative location within the irrigation network, and physical environment significantly shape agricultural adaptation, a number of dynamic factors such as local autonomy, effective monitoring, perceived risk, diversification potential, and community cohesion, and farmer experience explained much of the variation in cultivated extent observed across communities. Unlike the structural factors, these dynamic factors are relatively easy to influence and control. In Sri Lanka, increasing institutional support for the cultivation of other field crops could reduce water use in MASL systems and diversify the portfolio of options available to farmers during drought, though this support must be balanced with increased access to markets, market information, storage facilities, and agricultural inputs required to successfully cultivate these crops. Leveraging existing institutional structures to increase cross-scale communication between national and system-level water managers and farmers could increase information flow through the system and support system-wide adaptive capacity. Carefully planning infrastructural expansion to consider future population growth and shifting water demand could decrease the probability of future generations experiencing structural water scarcity. Officials should carefully monitor groundwater use to prevent overexploitation and to increase participation in collective cultivation activities. Finally, programs that support farmer responsibility and local resource control could be used to change farmer perceptions of risk and to increase water use efficiency.

CHAPTER 3

DYNAMICS OF INDIVIDUAL AND COLLECTIVE AGRICULTURAL ADAPTATION TO WATER SCARCITY

3.1. Introduction

In this chapter, we analyze survey and qualitative data to inform the construction of an agent-based model (ABM) that explores the role of water scarcity, farmer preferences, and technology diffusion in farmers' decisions to cultivate OFCs or paddy. Previous work on modeling farmer response to climate change and to potential water scarcity found that when climate forecasts are uncertain, farmers' response depends strongly on the way they think about risks and uncertainty (Hansen et al. 2004; Podesta et al. 2008). Much economic decision analysis and modeling of response to climate change assumes that actors will respond to risks and changes by making rational choices to maximize expected income or wealth, possibly with a degree of risk-aversion (Nordhaus 2008; Kolstad 2011), but a large body of empirical research in behavioral economics has found that people facing uncertainty decisions often use different heuristics to think about uncertainty and seek different objectives from simply maximizing expected wealth or income (Kahneman & Tversky 1979; Tversky & Kahneman 1992)

In analyzing the likely response of farmers to water scarcity, we drew on interviews conducted with key decision makers, water managers, and farmers during the 2013 and 2015 dry seasons as well as survey data collected in 607 households in twelve dry zone communities. Our qualitative data suggests that farmers are reluctant to cultivate OFCs for two reasons. The first is a strong cultural preference for paddy cultivation. Sri Lankan farmers have cultivated paddy for centuries and many government programs focus on supporting paddy cultivation. These include fertilizer subsidies, agricultural extension, and government purchase of paddy harvest at a set price (Jinapala et al. 2010). In addition, OFCs are difficult to store, so farmers must bring them to market immediately after harvest. In many cases, this causes market gluts at the end of the season that significantly reduce OFC prices for farmers. The second reason farmers cite for preferring paddy cultivation is the difficulty of cultivating OFCs during periods of extreme water scarcity. This may seem counterintuitive, since OFCs generally require less water than paddy, but the widespread practice of *bethma* drives many farmers to cultivate paddy when little water

is available. *Bethma* is an ancient practice in which farmers divide their fields and cultivate paddy on a subset of the command area. Under *bethma*, permanent field boundaries are temporarily abolished and land is redistributed amongst all farmers who cultivate in the command area. This redistribution process is complex and varies from system to system, but in general, each family received equal-sized parcels of land regardless of the amount of land owned (Spiertz & de Jong 1992; Thiruchelvam 2010b). During periods of extreme water scarcity, engaging in *bethma* is the best option for many farmers, as it ensures they are able to achieve modest yields for the season. In recent years, however, the diffusion of agrowells in the dry zone has allowed farmers to cultivate OFCs using groundwater during water scarce seasons (Kikuchi et al. 2001).

3.2. Fieldwork

We asked over 600 farmer heads of households whether they regularly planted OFCs in their irrigated fields. Responses y_i were labeled as 1 if farmers regularly cultivate OFCs in their irrigated fields and 0 if they regularly plant paddy in these fields, with $\Pr(y_i = 1) = \text{logit}^{-1}\beta_i X$. The respondent-level design matrix X is a set of binary indicators for key demographic variables including agrowell ownership, location in a high-capacity irrigation system, gender, ethnicity, land ownership, location at the head-end of a canal, and farmer organization membership. We also include a measure of socio-economic status constructed using household assets listed by interviewees. We computed the intra-class correlation (ICC) of the data to determine whether hierarchical modeling would be necessary. The ICC is essentially the ratio of between-group variance and total variance; it captures the proportion of the total variance in a value that is accounted for by grouping in the data. The ICC ranges from 0 to 1 where 0 indicates that grouping conveys no information and 1 indicates that all group members are identical (Gelman & Hill 2007). We find an ICC of 0.226, which suggests that clustering in the data explains a portion of the total variance in the dataset. Because of the hierarchical nature of the data, i.e. farmers nested within communities, and because of the strong community-level determinants of adaptive behavior identified in conversations with farmers and water managers, we group household survey responses by community to isolate variations in OFC cultivation that may be captured by community-level dynamics.

The data model characterizes the probability of a farmer i grouped within a community gn to engage in cultivation of OFCs in low-lying fields as follows:

$$\Pr(y_i = 1) = \text{logit}^{-1}(\alpha_0 + a_{gn[i]} + \sum_k \beta_k x_{ik})$$

where β is a vector of farmer-level coefficients, x is a vector of farmer-level covariates, α_0 is the intercept, a_{ij} accounts for state-level fixed effects where i is indicative of individual farmers and j represents a farmer's village. Farmer-level covariates are as follows: AW is a binary indicator of agrowell ownership, MAJOR is a binary indicator of location within a large surface water irrigation system, FEMALE indicates survey respondent sex, SINHALA indicates whether the respondent belongs to the dominant Sinhalese ethnic group or to a minority group, STATUS indicates high socio-economic status, LANDOWNER indicates whether the farmer is the legal owner of the land they cultivate, HE is a measure of the proportion of paddy fields cultivated by a farmer located at the head-end of their field canal, and FO indicates farmer membership in the local farmer organization.

Following Gelman 2008, we assign weakly informative Cauchy priors with a mean of zero and a standard deviation of 2.5 to each of the regression coefficients in the logistic regression. We tested to ensure these priors do not unduly constrain the posterior. Community-level variations not accounted for by the regression model were modeled using a normal distribution with a mean of zero and a standard deviation of σ_a , itself modeled with a Cauchy prior with a mean of zero and a standard deviation of 10. Formally: $\beta_k \sim \text{Cauchy}(0, 2.5)$

$$\alpha_0 \sim \text{Cauchy}(0, 10)$$

$$a_{gn} \sim N(0, \sigma_a)$$

$$\sigma_a \sim \text{Cauchy}(0, 10)$$

We used the `rstan` interface to the Stan Hamiltonian Monte Carlo software to perform the regression analysis. We ran each chain for 2000 iterations.

Figure 5: Predictors of OFC cultivation: Regression coefficients for choice to grow OFC instead of paddy. The dots show the median of the posterior probability distribution, the thick lines indicate the 66% highest-density interval and the thin lines indicate the 95% highest-density interval.

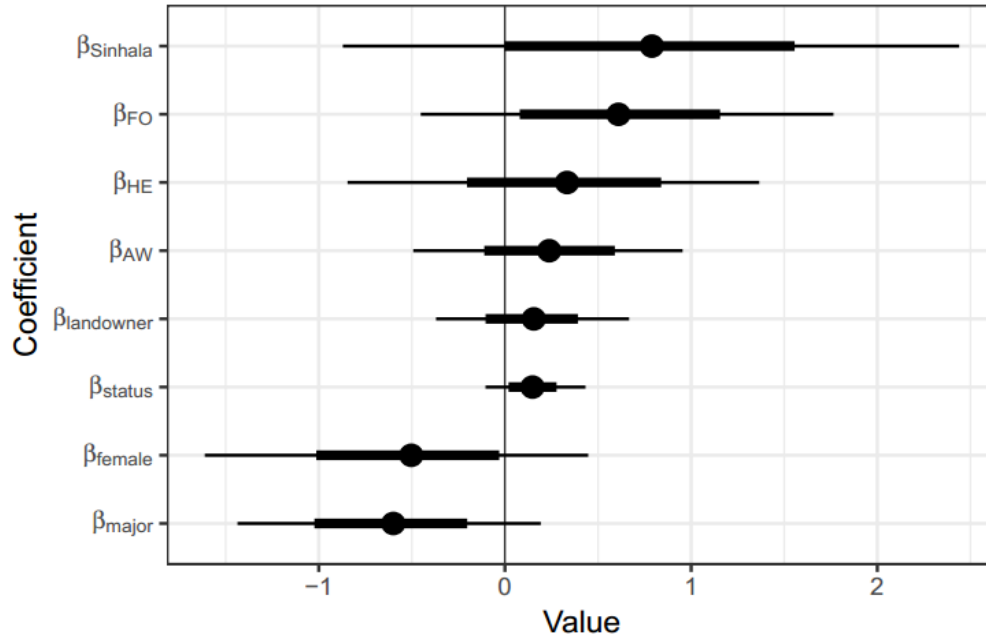


Figure 5 shows the results: though none of the effects are very strong, our results suggest that farmers who are Sinhalese and members of the local farmer organization are slightly more likely to plant OFC during the Yala season. Conversely, female farmers and farmers located within major irrigation systems are less likely to engage in crop diversification. The negative effect of location within major irrigation systems may capture the fact that to date the government has strongly promoted paddy cultivation within these large systems to ensure domestic self-sufficiency in rice production. There are strong, significant community-level effects for the majority of communities, suggesting that community-level factors drive much of the variation in OFC adoption.

3.3. Model design

We developed an agent-based model to study the role of farmer decision-making on adaptation to changing levels of water scarcity. We explore farmer adaptation across varying preference structures, forms of water access, and environmental settings. This model simulates a

single community of farmers, who share a distribution canal (DC) in the Sri Lankan dry zone. The DC is fed by a reservoir and distributes its water equally to a number of field canals (FCs), which carry water to the farmers' fields. Each farmer has a field on one FC. Collective action occurs at the FC level: each season, the farmers sharing an FC vote on whether that FC will collectively practice *bethma* for that season, and the decision follows the majority. This suggests that gradual changes in preference may produce abrupt effects when the number of supporters crosses the majority threshold.

Crop yields depend on access to water: the reservoir level, the amount of rain that falls on the DC (seasonal rainfall is uniform across the DC), and whether an individual farmer has an agrowell. Interviews with local water managers suggest that officials generally think about the level of water in a reservoir categorically (average, below average, or above average) rather than quantitatively, and choose seasonal operating and management strategies for the district accordingly. Similarly, farmers describe seasonal rainfall as wet, normal, or dry. Here, we present the model structure using the Overview, Design concepts and Details (ODD) protocol (Grimm et al. 2010). All code is available online at http://github.com/eburchfield/agrowell_abm.

3.3.1. Entities, state variables, and scales

The active entities in this model are farmers. Each farmer is characterized by an objective function, socioeconomic status, and agrowell ownership. The objective function characterizes how the farmer makes decisions under uncertainty. Following Podesta (2008), the possible objective functions are risk-averse expected utility, regret-adjusted expected utility, and prospect theory loss-aversion. These models are described in detail in Appendix A. Conceptually, these objective functions can be understood in the following manner. Risk-averse utility reflects a classical way of thinking about utility in human decision making; this objective function has emerged from the field of economics. Farmers using risk-averse utility to make decisions look across their portfolio of options for a season and seek to maximize profit. The utility of money declines as farmers' wealth increases, i.e. Rs. 100 has far more utility for a poor farmer than for a wealthy farmer. Regret-adjusted utility adjusts this classical way of thinking about utility for the regret a farmer may feel when considering some counterfactual situation, i.e. the farmer adjusts their utility for what might have happened had they made a different choice. Prospect theory emerged from research in experimental psychology (Kahneman & Tversky 1979). Under

prospect theory, decision makers anchor their perceived utility around a reference value. This could be a metric of peer performance, an individual's past performance, or some other value. Under prospect theory, the pain of losses relative to this reference point exceeds the pleasure in gains over the reference point.

Our logistic regression found that ethnicity was an important predictor of cropping decisions, but communities in the Sri Lankan dry zone are ethnically very homogeneous, so we did not include ethnicity in our model. There are 10 FCs on the DC, and each FC serves 15 farmers. There are no persistent state variables for DC and FC. Each iteration represents one growing season, and the simulations loop through 20 seasons.

3.3.2. Process overview and scheduling

At the beginning of the simulation, the farmers' state variables are initialized. At the beginning of each season (iteration), the level of water in the irrigation system's reservoir is randomly set to high (with 25% probability), medium (50%) or low (25%). Farmers know that there is a 25% probability of an especially wet season, a 50% probability of a "normal" season, and a 25% probability of an especially dry season. Based on these probabilities, farmers calculate expected utility for different crop choices (growing paddy under *bethma*, growing paddy without *bethma*, growing OFC with *bethma*, and growing OFC without *bethma*). The farmers then rank their preferences and the farmers of each field canal vote on whether to practice *bethma* in that season. After making the *bethma* decision, the farmers choose which crop to grow.

After the farmers make their cultivation decisions, seasonal rainfall is randomly set to wet, normal, or dry and the harvest yield is determined from a payoff table with some stochastic variance. The payoff table lists mean crop yields and variances for growing conditions, which were derived from government reports on crop prices and farmer self-reports of seasonal income reported in survey data. Finally, farmers' socioeconomic status is adjusted based on a balance of income and expenses. At this point, farmers who are sufficiently wealthy (socioeconomic status > 120,000 rupees, corresponding to one standard deviation above the population mean) will purchase an agrowell, financing it with payments over the following 10 seasons.

3.3.3. Design concepts

When farmers acquire agrowells, they gain an individual ability to irrigate their fields. This reduces their dependence on canal irrigation, which requires coordination and collective action among the farmers who share a canal. Under conditions of water scarcity, we hypothesize that expansion of agrowells may undermine traditional collective adaptations such as *bethma*. This is complicated because farmers must commit to planting before they know what the weather will be, and seasonal forecasts are very uncertain. Thus, the interaction between agrowells and collective action will be mediated by the details of how farmers make decisions under uncertainty.

We expect to see emergence occur through the collective decision-making about *bethma* on a field canal. If farmers' cultivation preferences change when they acquire agrowells, a gradual change in the number of agrowells could produce a sudden change in cultivation when a critical number of well-owners tips the balance in voting.

Farmers seek to maximize their objective function. Farmers compute their expected profit for each cultivation option, given known parameters (reservoir levels and agrowell ownership), under the probability distribution of seasonal rainfall (low, average, or high). When calculating prospect-theory utility, farmers use their income from the previous growing season as their reference point. Farmers calculate their objective function for the four possible cultivation decisions---choosing paddy vs. OFC and whether or not to practice *bethma* - and rank the choices from best to worst. They vote for or against *bethma*, with the majority ruling. After *bethma* has been decided for the field canal, each farmer then chooses between growing paddy or OFC. Farmers sense reservoir levels. They do not know what the weather for the upcoming season will be when they make their *bethma* and crop decisions. Farmers on the same field-canal interact by voting on whether to practice *bethma*. Farmers are initialized with random socioeconomic status. *Bethma* decisions are determined by the majority vote at the field-canal level, but after the field-canal makes this decision each farmer's crop choice is modeled as a Bernoulli process, using a logistic function to map the difference in expected utility between OFC and paddy onto a probability in the interval [0,1].

3.3.4. Initialization

Farmers are initialized with an agrowell ownership flag, socio-economic status, an objective function for decision making, and a set of risk parameters. Groups of fifteen farmers are randomly assigned to each field canal. Farmers' socio-economic status is drawn from a normal distribution with a mean of 100,000 rupees and a standard deviation of 2,000 Rs (Central Bank of Sri Lanka, 2014). A fraction of the farmers with a socio-economic status one standard deviation above the mean socio-economic status receive an agrowell at initialization. We assume that loans for these agrowells have been paid in full prior to initialization.

3.3.5. Submodels

Crop yields have a complex relationship with reservoir level, rainfall, whether farmers on a field canal practice *bethma* when the reservoir is low, and whether farmers growing OFCs have agrowells. We devised our crop-yield table based on survey data and records from the Sri Lanka Ministry of Agriculture. Paddy requires a great deal of water. If the reservoir level is normal or high, paddy will produce a full yield, averaging 100,000 Rs of seasonal profit, regardless of the rain. When the reservoir is normal or high and rainfall is normal or low (dry), OFC produces higher incomes than paddy, averaging 120,000 Rs; but in wet years (high rainfall), water damage to OFC and possible flooding reduce OFC yields to 90,000 Rs. If the reservoir is low, paddy yields will depend strongly on the amount of rain, producing 20,000 - 40,000 Rs. Practicing *bethma* can raise yields to 50,000 - 60,000 Rs. For all rainfall conditions, OFC produces 20% more income than growing paddy without *bethma*, but growing paddy with *bethma* produced higher yields for all levels of rainfall. Agrowells are especially valuable for OFC growers. With an agrowell, a farmer growing OFC can earn 84,000 - 144,000 Rs: much more than paddy under all conditions except high rainfall with a normal or high reservoir. This complex payoff table yields interesting dynamics under low-reservoir conditions. Normally, when reservoirs are low, it is economically advantageous for farmers to work together under *bethma*. This produces significantly higher income than either growing OFC or growing paddy without *bethma*. However, once agrowells enter the picture, those farmers who have agrowells can earn far more growing OFCs on their land, and thus they have an incentive to block *bethma*. Farmers with high socioeconomic status (more than one standard deviation above the population

mean) invest in agrowells. An agrowell costs 70,000 Rs, which is paid in annual installments over 10 seasons.

3.4. Results

We ran a 20-year simulation 100 times for each of four conditions of the farmer's objective function: all farmers using risk-averse expected utility, all farmers using regret-averse expected utility, all farmers using prospect theory, and a mixture with each farmer randomly assigned one of the three objective functions, with equal probability. Figure 6 shows how the fraction of field-canals choosing *bethma* varied with the penetration of agrowells.

Unsurprisingly, no field-canals choose *bethma* when the reservoir has an ample supply of water. But in conditions of water scarcity, the choice of *bethma* depends on the combination of the prevalence of agrowells. For all three objective functions, *bethma* drops to zero when a large fraction of farmers own agrowells.

Figure 6: Variation in *bethma* as a function of agrowell ownership for different reservoir levels and objective functions: The figures show the aggregate outcomes over 100 sequences of 20 growing seasons. Dots represent individual model runs and are jittered by 0.02 to aid visualization of overlapping points. The blue lines are lowess-curves.

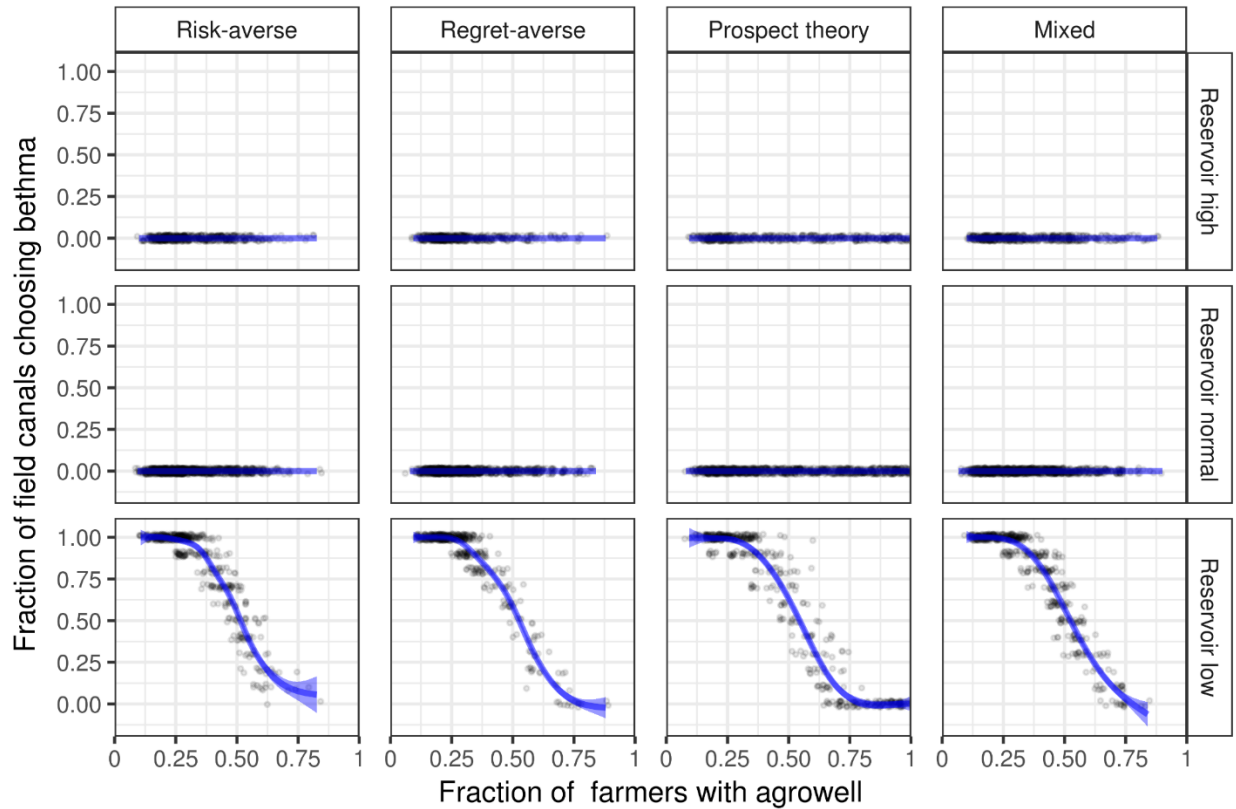


Figure 7: Variation in income over time for individual farmers, grouped by agrowell ownership, for different reservoir levels and different objective function: Each panel shows the simulations in which the reservoir is at a given level for a given season, so a simulation of 20 growing seasons will have dots that appear in the “Reservoir low” panel on those seasons in which the reservoir is low, in the “Reservoir normal” panel for those seasons in which the reservoir level is normal, and so forth. The dots represent individual farmers in 100 sequences of 20 growing seasons. The colored lines are lowess-curves fit to all the farmers in a panel with the corresponding agrowell-ownership. Where the dots form sets of three bands, the different bands correspond to different amounts of rainfall.

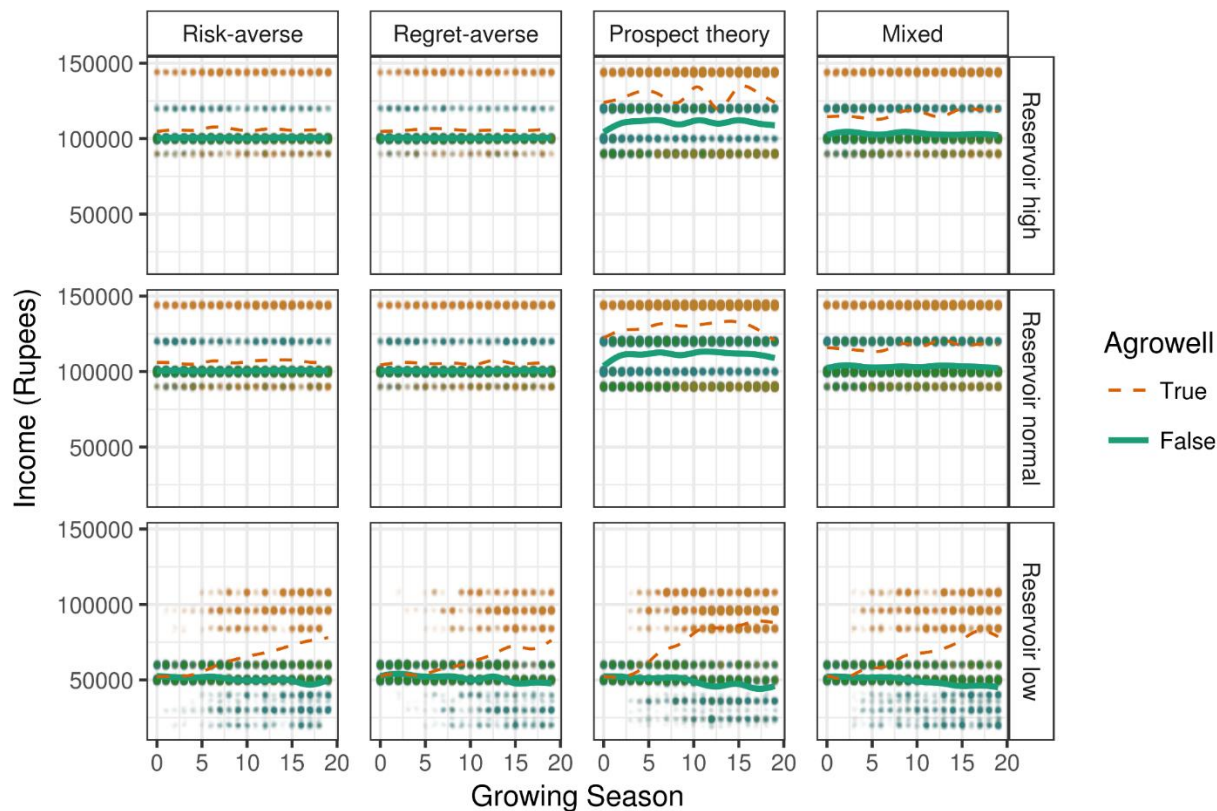


Figure 7 shows how the individual farmers' profits vary over time, broken down by the conditions of the reservoir. In general, farmers with agrowells have both greater average income and greater variation in income (because they are more likely to plant OFC when reservoir levels are normal or high, which makes them vulnerable to flooding and water damage if the rainfall is heavy that year). The decline of participation in *bethma* leads to growing income inequality in years with low reservoir levels because farmers with agrowells grow OFC without *bethma* and earn 80,000 - 108,000 Rs, depending on rainfall, but farmers without agrowells only earn 24,000 - 48,000 Rs, which is considerably less than the 50,000 - 60,000 Rs they would have earned growing paddy under *bethma*.

3.5. Discussion

The relationship between agrowells and *bethma* has complicated implications for policy. On the one hand, there is broad agreement among experts that farmers could be better off, both individually and collectively, if they would grow more OFC and less paddy. In particular, the lower water demands for OFC would relieve a good deal of stress on the water supply system

and make farmers more resilient to drought. Even when water is plentiful, agrowells can dramatically increase OFC yields. However, this simulation suggests that agrowells may also displace traditional collective responses to water scarcity, such as *bethma*. Agrowells are expensive and are thus out of reach for most farmers today. As successful farmers become wealthier and agrowells proliferate, tensions over *bethma* decisions may grow between farmers with agrowells and those without. In addition, as farmers with agrowells achieve majorities on field canals, farmers unable to afford agrowells may suffer economically, leading to growing inequality, as Figure 7 shows. However, the relationship between agrowells and *bethma* shows slight variations across different decision heuristics, so empirical studies of farmer views of risk and decisions under uncertainty could provide valuable information for policy analysts and decision-makers. This underscores the general observation by experts on risk and decision support that policies for managing environmental risks are more likely to be successful if they are grounded in empirical knowledge of people's actual behavior. It also highlights the importance of both environmental and social uncertainty in driving agricultural outcomes.

CHAPTER 4

FORECASTING VEGETATION HEALTH AT HIGH SPATIAL RESOLUTION

4.1. Introduction

Drought significantly reduces agricultural production, destabilizing food systems and threatening food security (Lesk et al. 2016). Remotely sensed measures of vegetation health, such as the Normalized Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI), are widely used to monitor spatiotemporal variations in the agricultural responses to drought (Rhee et al. 2010; Peters et al. 2002). Providing managers and farmers with accurate information about vegetation health increases system-wide capacity to prepare for and adapt to water scarcity (Dessai, 2009; Ziervogel et al., 2010). These indices can be used to identify vulnerable agricultural systems, to understand past agricultural responses to drought, and to guide efforts to increase resilience to future drought. Agricultural systems often exhibit nonlinear responses to sudden changes in water availability or human activity. However, many agricultural prediction tools rely on linear models to predict future vegetation health (Asoka & Vimal, 2015; Bolton & Friedl, 2013; Doraiswamy et al., 2005; Peters et al., 2002). Though more complex, nonlinear models have been used to predict rainfall in agricultural systems (Chattopadhyay & Chattopadhyay 2008; Singh & Borah 2013), metrics of agricultural drought such as vegetation health better capture changes in farmer livelihoods than the coarse resolution meteorological metrics of drought used in these studies. Coarse resolution models are not able to examine fine-grained intra-system dynamics and justify resource transfers. Higher resolution models tend to rely on datasets that are only available in data-rich regions of the world (Koide et al. 2012; Mo et al. 2005; Bolton & Friedl 2013; Kogan et al. 2012). Furthermore, data scarce regions tend to lack the economic resources required to buffer against the effects of drought.

Our objective was to create a user-friendly predictive tool that will increase the capacity of data-scarce agricultural systems to prepare for and respond to drought in the future. We have created a tool that (1) predicts future vegetation health values at a (2) high spatial resolution using (3) open source tools and data that are (4) global in coverage. With simple user inputs, our tool downloads, processes, models, and forecasts vegetation health at 16-day intervals at a 250 meter resolution anywhere in the world. The tool applies a gradient-boosted machine model to

Moderate Resolution Imaging Spectroradiometer (MODIS) datasets openly available on NASA's LP DAAC server. The model learns potentially complex relationships between past remotely sensed variables (and their interactions) and future vegetation health as measured by the Enhanced Vegetation Index (EVI). In this paper, we apply the tool in two locations: Sri Lanka and California.

4.2. Experimental Design: Location and Data Type

We designed an experiment across location and data dimensions to assess how well our process performs under different conditions. In terms of location, we hypothesized that within each data category, the model would perform better in California, where there are fewer clouds than in tropical Sri Lanka. We anticipated that the spectral data (MOD09A1) would predict vegetation health better than Level 3 MODIS data products (land surface temperature, leaf area index, etc.), which are derived from the spectral data, because the flexible models will, in effect, learn intermediate representations of the underlying data that are more suited to predicting future EVI values than the NASA-derived representations of that same underlying spectral data. Note that while the MOD09A1 datasets are less processed than the Level-3 products, these datasets have undergone significant processing. These products are estimates of spectral surface reflectance in each band of light as if there had been no scattering or absorption caused by the atmosphere. The “raw” data collected by the MODIS satellite undergoes correction for atmospheric gases, aerosols, clouds, and water vapor. The highest quality observation over an 8-day period of time is extracted as the observation in the final 8-day resolution MOD09A1 dataset (EROS Data Center, 2017).

From a machine learning perspective, the Level 3 products are part of a feature engineering process orthogonal to the learning task of mapping spectral data to future EVI. We hypothesized that models with only lagged EVI as a predictor will have the lowest performance because all the other predictor sets are multivariate supersets, containing the underlying data from which lagged EVI is computed and more. If the additional variables added little predictive power, we anticipated that the model would learn to ignore them. We included this univariate lagged EVI model to measure relative prediction error reductions associated with land use and time, Level 3, and spectral data. Similarly, because land use and time are included in both the

Level 3 and spectral models, we tested models including lagged EVI, land use classification and time of the year.

4.2.1. Experimental Variable: Location

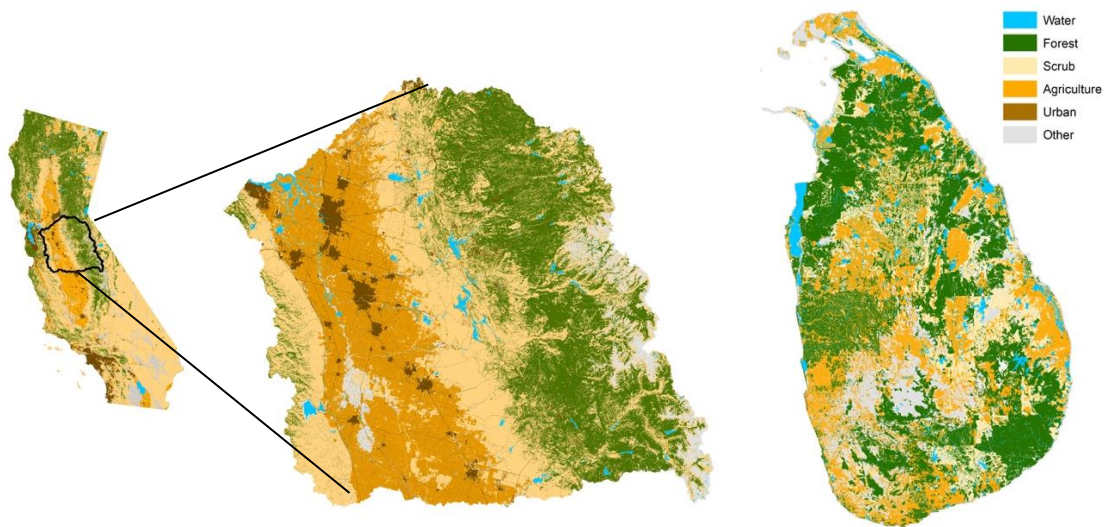
We selected two regions with distinct agroecologies, climates, and levels of data availability: Sri Lanka and the San Joaquin Valley in California. Sri Lanka is a small island nation located off of the eastern coast of India that covers approximately 66,000 square kilometers and is home to nearly 21 million people (Government of Sri Lanka 2010). The country receives rainfall during two monsoon periods. The northeast monsoon lasts from October to December and brings two-thirds of annual rainfall to Sri Lanka. The southwest monsoon lasts from May to October and brings rain primarily to the southwestern region of the island. This rainfall pattern divides the island into wet and dry zones and creates two distinct cultivation seasons, the wet *Maha* season and the dry *Yala* season (Samad 2005; Senaratne & Scarborough 2011). During the wet season, most farmers cultivate rice. Rice is a staple of the Sri Lankan diet and an estimated 30 percent of the total labor force is involved in rice production (Mahaweli Authority of Sri Lanka 2012). Farmers capture wet season rainfall in reservoirs and cultivate rice during the dry season with stored water. During water scarce dry seasons, farmers cultivate other field crops such as soy, maize, and grain. Increasing numbers of dry zone farmers pump groundwater to irrigate other field crops (Kikuchi et al. 2001). Field size is small in Sri Lanka, with over 70 percent of farmers cultivating less than 2.5 acres of land (Withananachchi et al., 2014). Persistent cloud cover year-round significantly reduces remotely sensed data availability.

The San Joaquin Valley in California covers approximately 40,000 square kilometers and is home to over 1.6 million people (California Department of Water Resources, 2013). This valley is one of the most productive agricultural systems in the world, with an annual gross production of more than 25 billion dollars (EPA, 2013). The average farm size is 162 acres, significantly larger than the small plots held by Sri Lankan farmers (California Department of Water Resources, 2013). The primary crops cultivated in the area are grapes, walnuts, almonds, and cherries (California Department of Food and Agriculture, 2013). As in Sri Lanka, many of the agricultural fields in the valley receive water from surface water irrigation systems. Heavy groundwater pumping also provides a significant amount of agricultural water in the region

(California Department of Water Resources, 2013). The climate in the valley is Mediterranean, with moderate temperatures throughout the year. Cloud cover is significantly lower than in tropical Sri Lanka.

These two regions were selected for the following reasons. First, in both regions, irrigation infrastructures allow decision-makers to move large amounts of water over considerable distances. Decision-makers may have the capacity to respond to our predictions by moving water to areas we predict to have relatively low vegetation health. Second, the differences in agricultural field size and crops cultivated tests the performance of our models in regions with markedly different agroecological systems (Figure 8). Finally, by comparing model performance in the cloudy tropics and relatively cloud-free California, we can analyze the effect of data availability over a fixed time interval (11 years) on predictive performance.

Figure 8: Land use in the San Joaquin Valley and Sri Lanka.



4.2.2. Experimental Variable: Data Type

Remotely sensed measures of vegetation conditions have been used in many studies to monitor the agricultural effects of drought (Brown et al., 2002; Ji, L., Peters, 2004; Thenkabail, Gamage, & Smakhtin, 2004). We measured these effects using the Enhanced Vegetation Index (EVI) which is a proxy for the health of agricultural crops (Cai & Sharma 2010; Gumma 2011; Sakamoto et al. 2005; Xiao et al. 2006; Galford et al. 2008), highly correlated with the leaf area

index (Huete et al. 2002; Sakamoto et al. 2005) and positively linearly related to vegetation fraction estimates (Small & Milesi 2013). The EVI is measured as:

$$EVI = G \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + C_1 \times \rho_{RED} - C_2 \times \rho_{BLUE} + L}$$

where ρ is atmospherically corrected surface reflectance, L is the canopy background adjustment, and C_1 and C_2 are the coefficients of the aerosol resistance term, which uses the blue band to correct for aerosols in the red band (Huete et al. 2002). EVI values approaching one indicate high levels of photosynthetic activity. For predicting EVI, our analysis compares the performance of four sets of predictor variables:

- (1) Land use, time period, the value of EVI from the last time period, and spectral data from the previous time period,
- (2) Land use, time period, the value of EVI from the last time period, and Level-3 MODIS products (land surface temperature, NDVI, leaf area index, the fraction of photosynthetically active radiation, net photosynthesis, and gross primary productivity) from the previous time period,
- (3) Land use, time period, and the value of EVI from the previous time period, and
- (4) The value of EVI from the previous time period.

We included the third and fourth options because simple univariate models leveraging past values of a variable are often effective in forecasting future values of the same variable, especially if those values are adjusted for the time period (seasonal effects). For options 1 and 2, we also included the lagged population and El Nino sea surface temperature index.

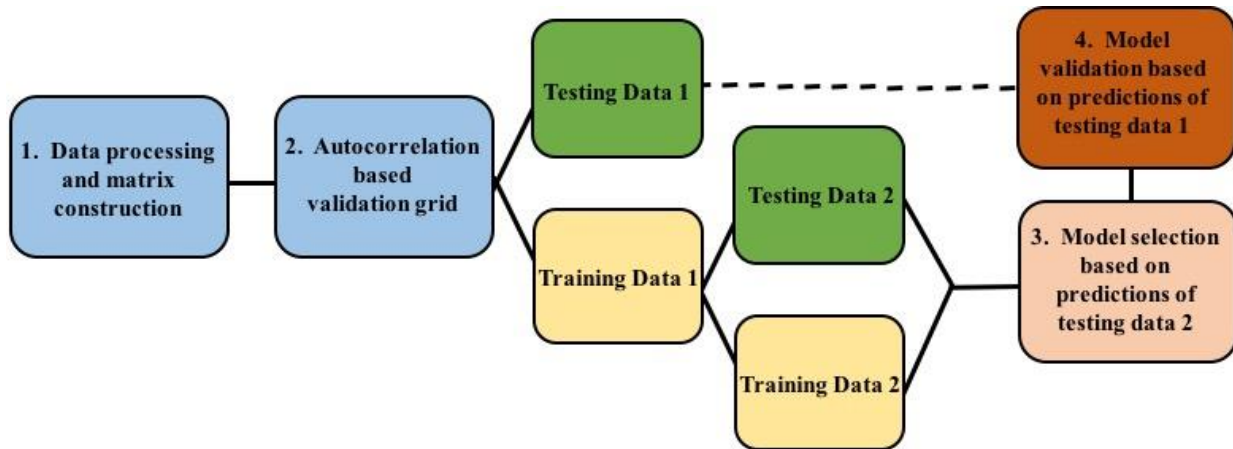
Table 1: Description of the datasets used in the predictor sets.

	MODIS product	Layer	Description
Spectral Model	MOD09A1.005	B1_lag	Lag of MOD09 band 1, 620-670 nm
		B2_lag	Lag of MOD09 band 2, 841-876 nm
		B3_lag	Lag of MOD09 band 3, 459-479 nm
		B4_lag	Lag of MOD09 band 4, 545-565 nm
		B5_lag	Lag of MOD09 band 5, 1230-1250 nm
		B6_lag	Lag of MOD09 band 6, 1628-1652 nm
		B7_lag	Lag of MOD09 band 7, 2105-2155 nm
Level 3 Model	MOD11A2.005	LST_Day_1km_lag	Lag of daytime land surface temperature
		QC_Day_lag	Lag of quality control for daytime LST
	MOD13Q1.005	EVI_lag	Lag of enhanced vegetation index
		NDVI_lag	Lag of normalized difference vegetation index
MOD15A2.005	VI_Quality_lag	Lag of quality control for vegetation indices	
	Fpar_1km_lag	Lag of fraction of photosynthetically active radiation	
	Lai_1km_lag	Lag of leaf area index	
MOD17A2.005	Fpar_Lai_QC_lag	Lag of quality control for FPAR and LAI	
	GPP_lag	Lag of gross primary productivity	
Level 3 & Spectral	Ancillary data	PSN_lag	Lag of net photosynthesis
		Land_use	SL Survey Department, National Land Cover Database
		nino_lag	Lag of El Nino sea surface temperature index
		GWP_lag	Lag of population

4.3. Methods

We downloaded and processed eleven years of remotely sensed imagery (2004 – 2014). We combined this data with ancillary datasets and reshaped it into a single matrix where each row corresponds to a pixel at one time and each column is a measured variable. We divided the observations into Training Data 1 and Testing Data 1 by sampling from large spatial grid indices without replacement (Figure 9). We then divided Training Data 1 into Training Data 2 and Testing Data 2 with the same spatial sampling process, and trained multiple models on Training Data 2, varying the hyper-parameters for each model estimation. We used Testing Data 2 to assess the performance of each model’s predictions. We repeated this loop of learning on Training Data 2 and testing on Testing Data 2 for each of the four different data types, and chose the combination of data type and hyper-parameter setting that achieved the highest performance in predicting Testing Data 2. Finally, we validated the best-performing model from the previous step by testing its performance on the held-out data in Testing Data 1. We repeated this entire process separately for Sri Lanka and California. This process is summarized in Figure 2 and detailed in the next subsections.

Figure 9: Methods overview.



4.3.1. Data Processing and Matrix Construction

We created a set of Python scripts to automate downloading and processing MODIS data from the MOD09A1, MOD11A2, MOD13Q1, MOD15A2 and MOD17A2 datasets. These scripts, along with the modeling and validation scripts, are open source and can be used on any MODIS dataset found on NASA’s LP DAAC server, for any region of the world in which MODIS data is collected. The required user inputs include the MODIS tiles for the region of interest, the first and last download dates, and the path to a reference image. The reference image stores information about the desired projection and extent final dataset. Users unfamiliar with Python can create the reference image using the MODIS Reprojection Tool or user-friendly software such as ArcGIS. The user has the option of including ancillary geospatial datasets such as land use information, socioeconomic data, or climate data. For our analysis, we included gridded world population (CIESIN, 2005), land use (Survey Department, 2011; Homer et al., 2015) and an El Niño sea surface temperature index (Rayner et al. 2003). The Niño 3.4 SST Index was used in Sri Lanka and the Niño 4 SST Index in California. The software downloads, mosaics, clips, and projects HDF files downloaded from the LP DAAC server and masks all pixels not flagged as “good quality” by each dataset’s quality mask. In both locations, particularly in Sri Lanka, this created a large amount of missing data. 8-day datasets are transformed to a 16-day time step by computing the average of two quality-masked 8-day pixels. All datasets are resampled to match the spatial resolution of the EVI dataset (250 meters). The software reshapes the stack of images for each dataset into a single column and stacks columns

to create a two-dimensional matrix with dimensions pixel-time by number of variables. The software also creates columns describing the time period of each observation (dividing the year into 16-day periods), the latitude and longitude of each pixel, and the pixel's location in the autocorrelation-based validation grid.

4.3.2. Autocorrelation-based Validation Grid

We computed the spatial autocorrelation functions of the MOD13Q1 imagery to divide the final matrix into a grid of independent areas. In the case of both the San Joaquin Valley and Sri Lanka, the autocorrelation functions approached zero at a lag of 150 pixels (approximately 35 kilometers). We constructed a grid of 150-pixel by 150-pixel cells, each with a unique identifier. A random subset of these cells were selected as training data and the remainder were used as testing data. This reduces spatial autocorrelation between our testing and training datasets to allow performance of the model on the testing data to estimate how well the model will predict new data that is collected after a model is trained.

4.3.3. Model Training and Selection

We selected two model types that have consistently performed well in supervised learning tasks with large amounts of training data where potentially complex functions link the predictor and outcome variables: gradient boosted machines (GBM) and deep neural networks. To contextualize quantitative performance measures of our models (correlation and mean-squared error between vectors of predicted and actual EVI), we compared them to a baseline model that serves as a proxy for potentially currently available forecasting undertaken by local residents. Ideally, our baseline model would be a univariate time series model fit to the training data that uses past values of EVI in the hold-out data to forecast future EVI, a standard model for time series forecasting in the environmental sciences, but due to very large amounts of missing EVI in Sri Lanka this was not feasible. There are often gaps between observed values of EVI for many consecutive time periods due to cloud contamination. To approximate the desired baseline model, we created a simple model that uses approximate nearest neighbor search to search for k pixel-time observations approximately closest in space and time in the hold-out data (with the condition that the time is in the past) and averages their values of vegetation health to predict the hold-out data EVI. If the search does not return any neighbors because no neighbors without

missing EVI data can be found within the k results, the algorithm uses the average of all EVI values up to that point in time as the prediction.

The GBM combines gradient-based optimization, which iteratively adjusts model parameters in the direction of lower training data prediction errors by using gradient computations, and boosting, which improves an ensemble of weaker base models by adjusting the training data. The base models are trees that divide predictor variable values into distinct regions by choosing variables to make binary splits on, and the threshold values of those variables where the split should be made (Hastie et al. 2009). An important desideratum for our modeling algorithms was automatic handling of missing predictor variable values. Remotely sensed datasets used to detect vegetation health often have many missing values due to cloud cover. The GBM can handle missing predictor variables by incorporating them in the overall tree structure by always moving missing values to the left at splits in the trees. Furthermore, the model does not rely on one-hot-encoding of categorical variables so our time and land-use factor variables, which have many levels, are handled efficiently.

Using large trees allowed the model to automatically learn higher-order interactions between predictor variables. Although our largest models only had slightly more than 10 predictor variables, interactions between variables, e.g. lagged EVI and lagged Band 7, may improve predictive power. The level of interaction to which the model may search depends partly on a hyper-parameter that we tuned on the training data (see next paragraph). If we were using a linear regression model, interactions between variables would need to be specified manually; however, manually specifying all such potential interactions would be prohibitively time-intensive. Furthermore, the exact interactions that lead to the best predictive performance likely vary by location and thus would need to be specified by local experts each time the model was applied to a new location. The GBM algorithm implicitly automatically tries many interactions and learns which are useful from the data.

There are three important hyper-parameters for the GBM that need to be set for the model to be estimated. They can affect overall model complexity and thus whether the model over-fits training data or generalizes well to new data, but their best values depend on the nature of the data and the prediction task and can rarely be effectively determined a priori. It is common to conduct an exhaustive grid-search over the entire (suitably discretized) hyper-parameter space. In

fact, this is the only automated option available in most statistical software. However, this may be too slow for data this large unless the user has access to a large cluster of powerful computers. It is more efficient to use past model runs to learn better values. Given some prior distributions that cover most of the possible values of the hyper-parameters, we used the data to adaptively learn a meta-model of the effect of the hyper-parameters on the mean-squared error of the model's predictions of Testing Data 2, and use that meta-model to search through the hyper-parameter space for an appropriate setting of the hyper-parameters using a Tree of Parzen Estimators search algorithm (Bergstra et al. 2013). This online learning of hyper-parameters automates the entire model building process. The user is not required to specify anything other than the location in the world, after which the scripts download the data and train a model specific to that location.

We also trained feed-forward deep neural networks. However, we abandoned this model in the further modeling experiments after observing consistently poorer predictive performance on Testing Data 2 in Sri Lanka and California and significantly longer training times, which included a process we developed for imputing missing predictor variable data because this model cannot automatically handle missing data. The lower performance may be due to the larger number of hyper-parameters that need to be tuned. It is likely that if we could devote more computing resources to hyper-parameter search that the deep learning algorithm could learn a competitive model. Our objective, however, was to develop a tool that could be used with reasonable computational resources (i.e. a single powerful computer, not a cluster of computers).

Our model selection process involved selecting (1) the best performing model type (GBM or deep learning), (2) the hyper-parameter values for that model type, and (3) the set of predictor variables (Spectral Data, Level 3 Products, land use and time period, or lagged EVI). By “model selection,” we mean a specification of all three components.

4.3.4. Model Validation

We trained models on Training Data 2 and selected the model that performed the best on Testing Data 2. Then we trained the model with those hyper-parameter settings and data type on the full training data, Training Data 1. Finally, we used this model to forecast all the 16-day-ahead values of EVI in Testing Data 1, the hold-out data. We used a flexible model that can

learn complex relationships, such as the interactions discussed above. However, if the model is not tested on data separate from the data it was trained on, there is a risk that the model may have learned structure that is unique to the training data and not generalizable to the ultimate task of predicting EVI for new observations. Although we only used Testing Data 2 for tuning the three hyper-parameters of the GBM and selecting which set of predictor variables is most effective, there was still a risk that we may have over-fit Training Data 1 (Training Data 2 and Testing Data 2) and learned characteristics of the noise in this data in addition to the characteristics of the signal. Therefore, to test our best model on fresh, unseen examples, the model predicted the observations in Testing Data 1, which was only used for this purpose.

4.4. Results

4.4.1. Model Selection

Figure 10: Model performance for each data type in California and Sri Lanka as measured by the percent reduction in mean squared error below the lagged EVI model for each location.

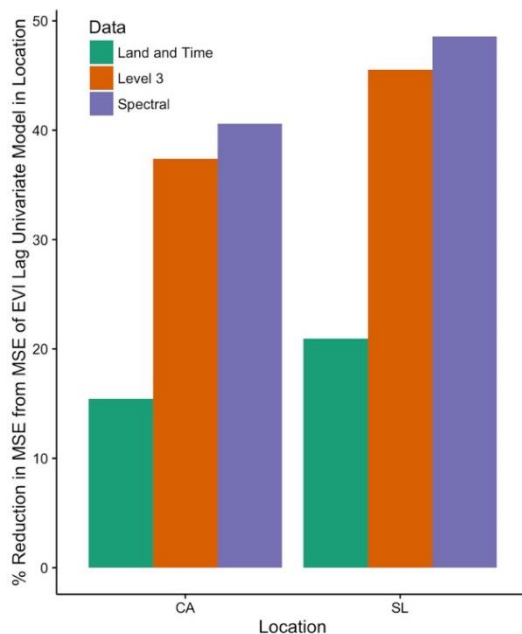


Figure 10 plots the percent reduction in mean squared error (MSE) below the MSE of the GBM model using only lagged EVI. We plot the results for the best hyper-parameter setting found for each experiment after using the same model construction and selection scripts for the

eight possibilities (four data types and two locations). In both locations, when we included additional datasets the model learned useful relationships between these datasets and future EVI and error dropped compared to the simple lagged EVI model. All three data types in the plot also used lagged EVI as a predictor variable, and the Level 3 and Spectral data types used land use and time as predictors, allowing us to determine the relative importance of adding additional data.

Although the absolute performance of models varies across locations with different levels of data availability and agroecology (which we explore in the next section), in both locations the magnitude of error reduction between the predictor variable sets is similar. For instance, error is reduced by between 40 and 50 percent when moving from only lagged EVI to lagged EVI and Spectral data. Overall, these results accord with what we anticipated (see Table 1): the predictor Data ordered by performance is Spectral, Level 3, land use and time, lagged EVI, and performance is higher in California. We used these results to select the spectral data for both locations and estimated the model with the chosen best performing set of hyper-parameter values on the full Training Data 1. Finally, we used the estimated models to make predictions on the held-out data in both locations to validate our model and compare to a baseline.

4.4.2. Holdout Validation

We measured the performance of the model by calculating the correlation between the vector of 16-day ahead predictions of EVI and vector of actual values of EVI in the held-out data. We computed the correlation for each land use category and found that model performance relative to the baseline is high in all categories of land use (Figure 11). Performance in California is higher because of more cloud-free days and less missing data. In both regions, the correlation in agricultural areas is above 0.75 (0.86 in California and 0.76 in Sri Lanka). Predictive power more than doubles in agricultural areas compared to the baseline model.

In Figure 12 we plot the performance for held-out agricultural pixels. The x-axis histogram displays the distribution of hold-out *predicted* agricultural EVI values, and the y-axis displays the distribution of *actual* agricultural EVI values. If our model made perfect predictions, all points in the scatter plot would line up on the dotted line. In Sri Lanka, the strongest predictions of EVI are at values indicative of healthy vegetation, between 0.5 and 0.8.

Predictive performance decreases for low EVI values, which are suggestive of stressed vegetation or atmospheric noise. The low predictive performance for extreme EVI values in Sri Lanka may be due to high levels of atmospheric noise. In California, the drop in performance for low EVI values is very slight.

Figure 11: Correlation between Predicted and Actual EVI in California (A., n=61,681,296) and Sri Lanka (B., n=36,831,863).

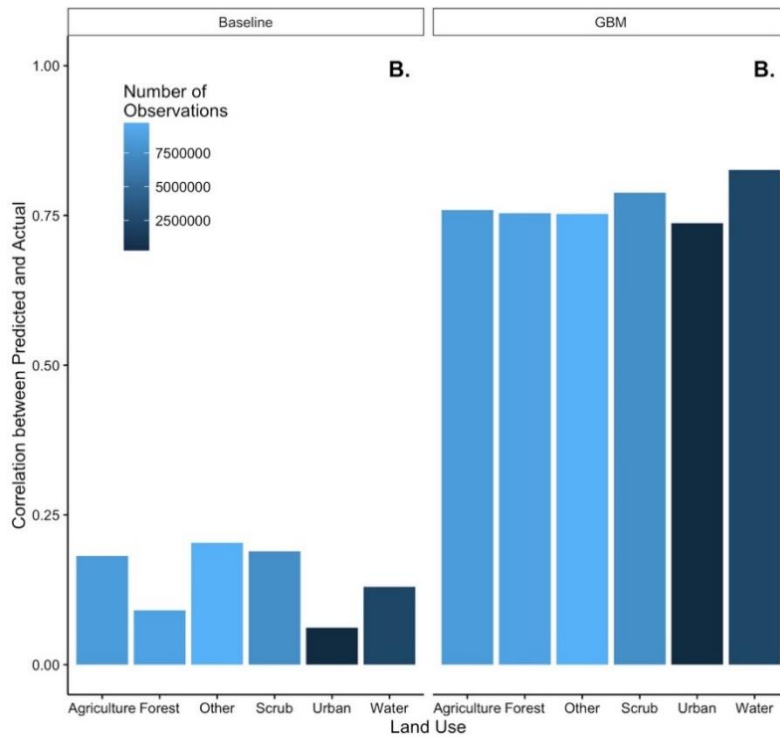
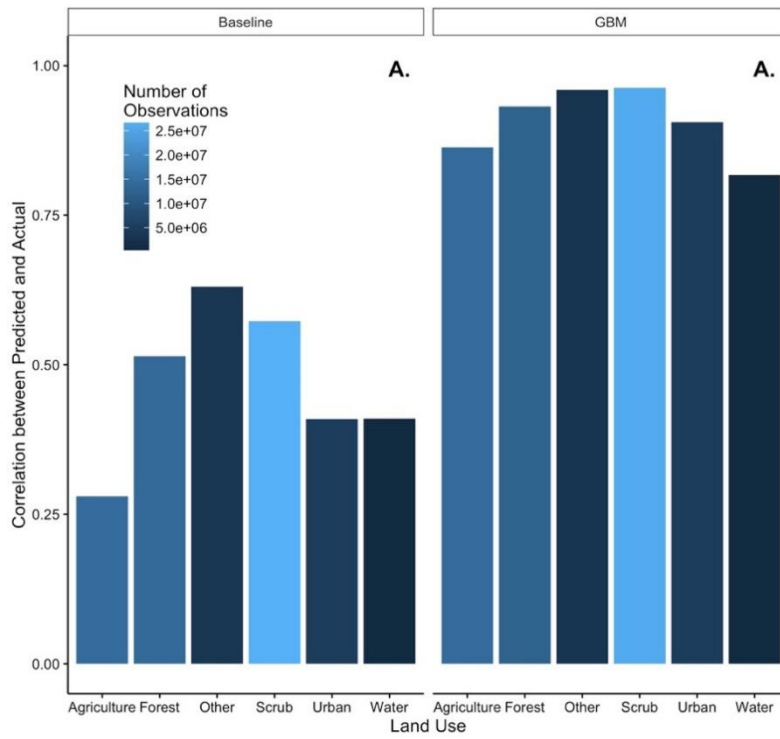


Figure 12: Performance across values of true measured EVI in California agricultural land (A., n=14,414,402) and in Sri Lanka agricultural land (B., n=8,402,076).

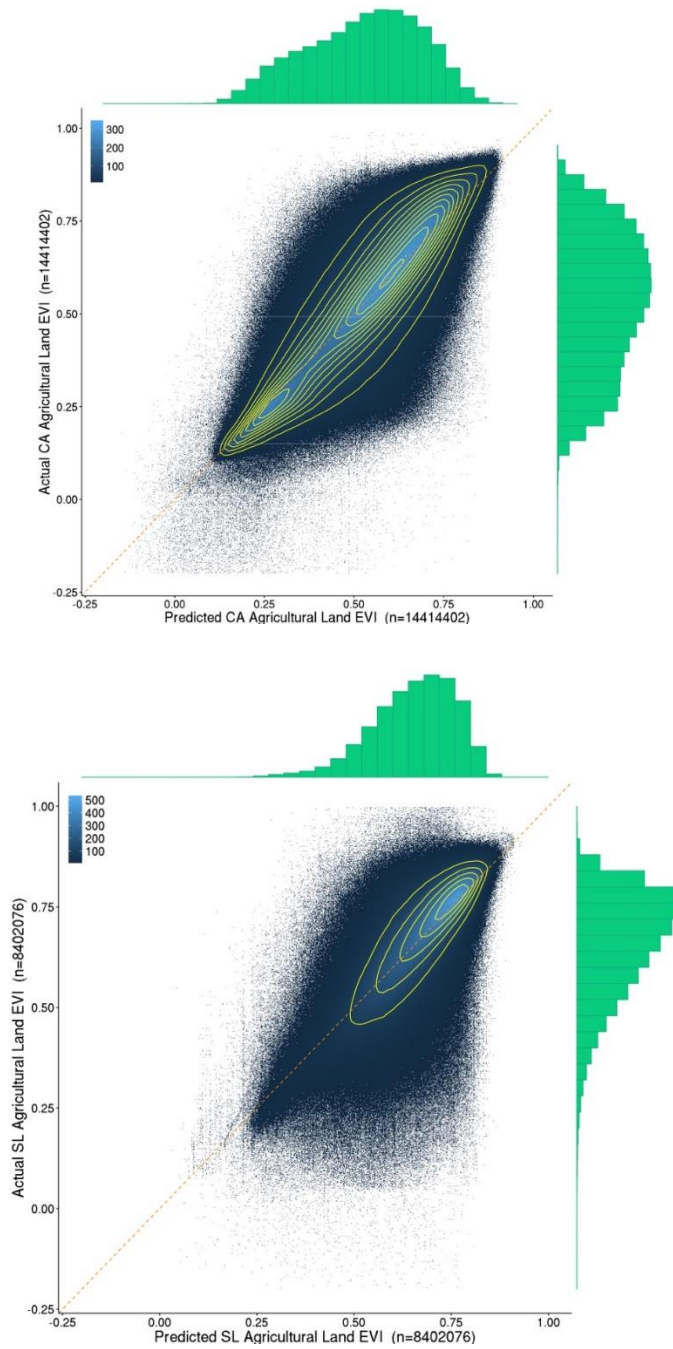
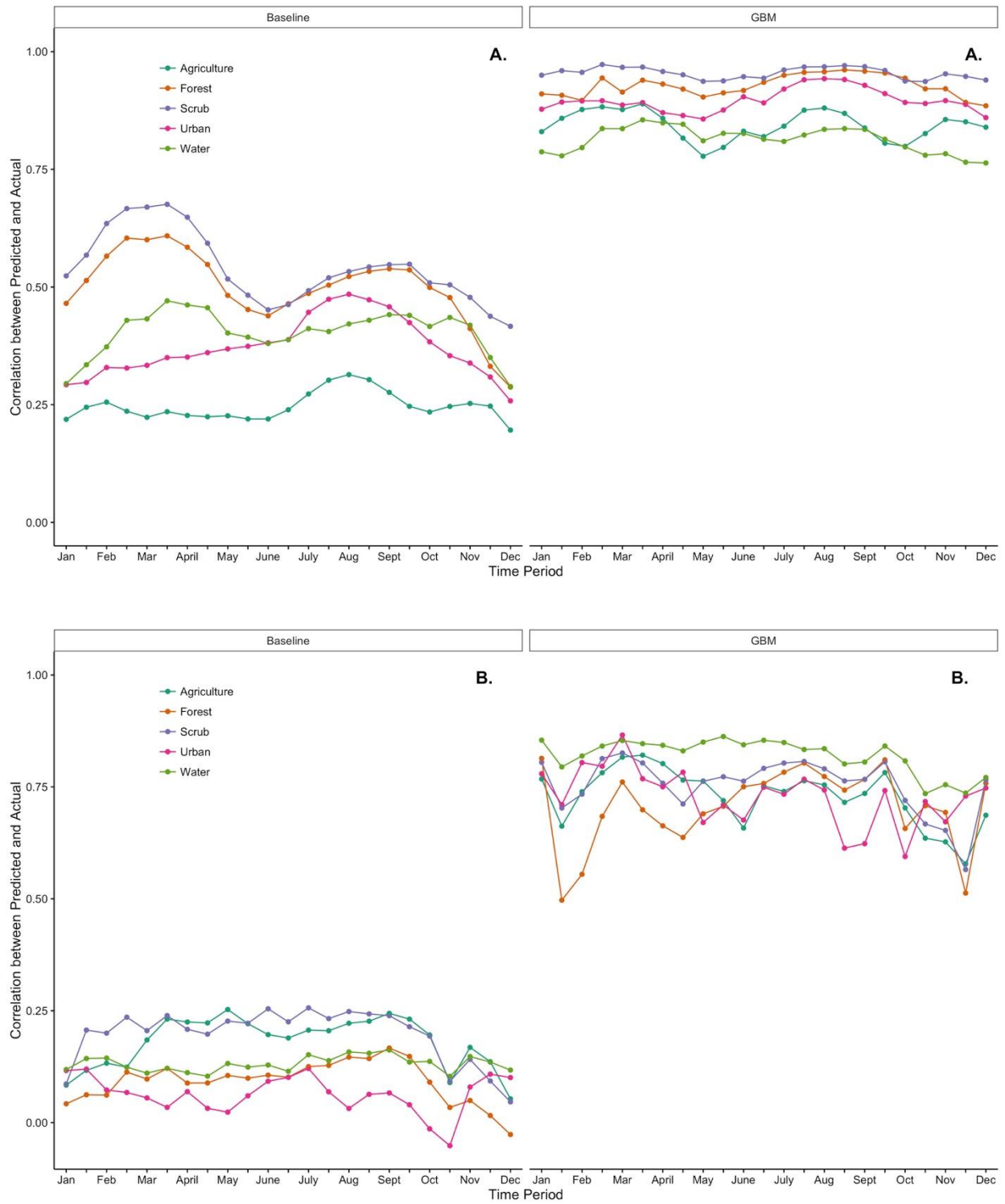
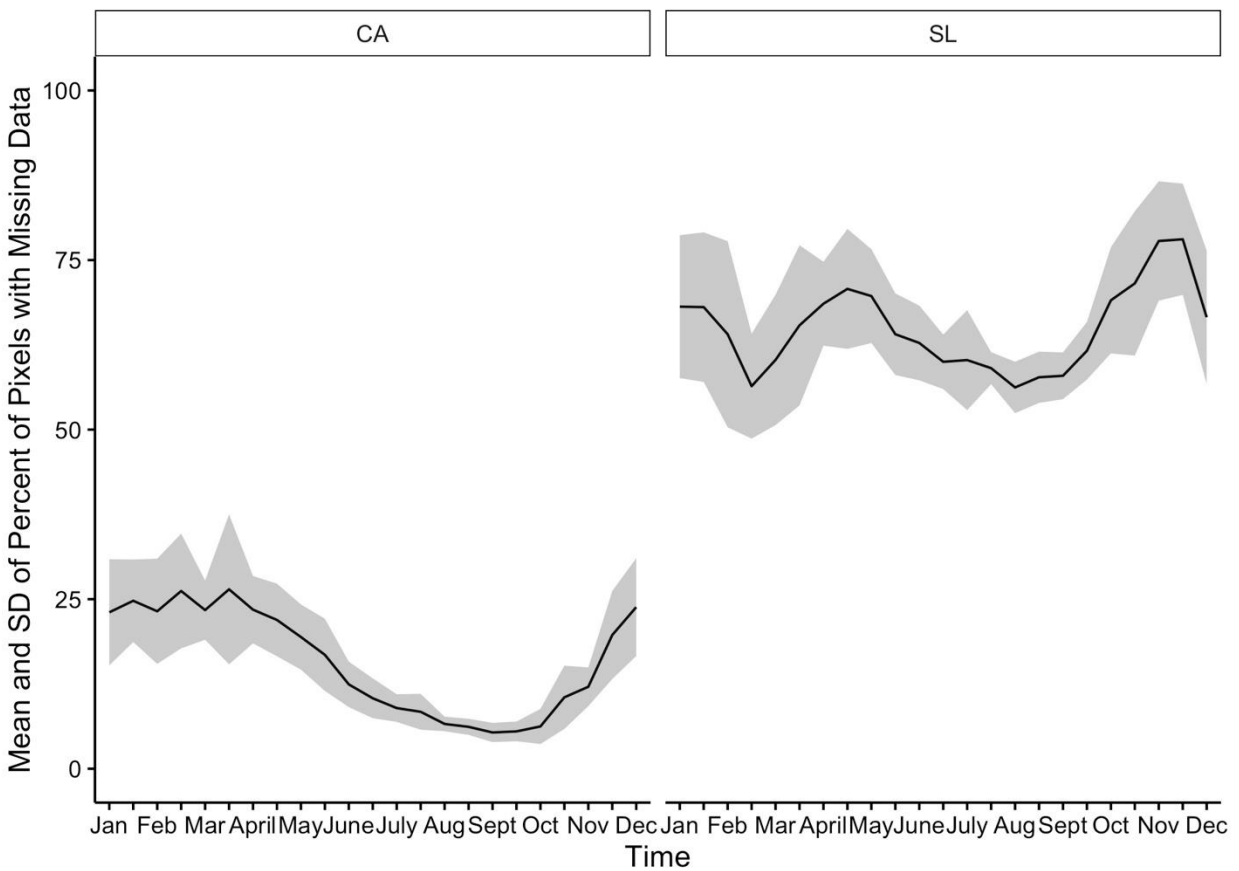


Figure 13: Correlation between predicted and actual EVI over time periods in California (A.) and Sri Lanka (B.).



In Sri Lanka, there was variation in the performance of our model across periods of the year (Figure 13). We plotted the average percent of missing data at each time period of the year (Figure 14) and found that the drops in correlation occurred after increases in the percent of missing data. Many of the lowest drops in correlation occurred during the *Maha* wet season (October – February), during which the majority of the island is covered in clouds. In California, the performance of the model is consistently high across land use categories and time periods. Periods of lower correlation occur during the winter, when there is also the highest extent of masked data.

Figure 14: Percent of pixels with missing data over 23 16-day periods of the year.



4.5. Discussion

Agricultural communities around the world are experiencing increased climate unpredictability. Scientists have built models to monitor and predict changes in agricultural health, but many of these models fall short in one of four ways. First, many models rely on

proprietary software or data and fail to publish fully reproducible results and software. Second, high resolution analyses are often only undertaken in specific regions due to data constraints. Third, few analyses are global in coverage. Finally, many existing analyses focus on describing and explaining processes rather than forecasting. Models that do forecast are not often rigorously tested out of sample on held-out data.

We have addressed these shortcomings in this paper by designing and testing a user-friendly set of scripts that download, process and predict high resolution values of vegetation health for any MODIS tile. Our tool makes predictions at a 250-meter resolution, which captures field-level variations in vegetation health and can support local and regional decision-making. All scripts and data are open source (<http://johnjnay.com/forecastVeg/>) and well-documented (see the Supplementary Materials). The tools we have constructed can be applied to any region in which MODIS data is collected. While this tool is best suited for regions with low cloud cover, it still performs well in one of the cloudiest regions of the world. Finally, our model is tested on held-out data which increases the likelihood that it will perform well in practice, and has high predictive power across land use categories and throughout time periods. The tool can be used to monitor and predict vegetation health at a high resolution in regions in which no local data is available, where it could support agricultural decision-making.

Though the scripts were designed for the prediction of EVI, they can be used in a number of ways. The data download and processing scripts that generate the input for the GBM model allow users to create large spatiotemporal data-cubes of any MODIS dataset with a simple one-line command. These datasets can be used to explore past trends in vegetation, investigate the effects of environmental stressors such as droughts and floods on vegetation health, and monitor inequalities in water access across space and time. The option to include high-resolution local ancillary datasets could significantly increase the predictive power of the models. Future research could combine our scripts with additional ancillary data to model the effects of particular social and institutional factors on vegetation health. In addition, the integration of supervised machine learning techniques and remote sensing could be used to model human-environmental interactions and predict other environmental phenomena.

CHAPTER 5

CONCLUSION

The demographic, economic, and environmental changes facing Sri Lanka challenge agrohydrological systems around the world. Research exploring how these complex resource management systems respond to water stress is of paramount importance if we are to meet growing demands in an increasingly stressed physical environment. The research presented in this dissertation is a step towards understanding the factors that drive agricultural adaptation and ultimately increasing the adaptive capacity of agricultural communities in the Global South. This research highlights the importance of local and dynamics factors in driving agricultural adaptation. My research in Sri Lanka reveals the importance of local autonomy, effective monitoring of resource use, perceived risk, diversification potential, community cohesion, and farmer experience. Unlike physical environment, climate, and infrastructural capacity, these factors are relatively easy to influence and control. Through agent-based modeling, this research reveals the tradeoffs between collective and individual agricultural adaptation. We show that the proliferation of an individual adaptive strategy, agrowell use, may displace traditional collective responses to water scarcity, namely the practice of *bethma*. This displacement may increase income inequality in agricultural communities, exacerbating the vulnerability of those with few resources. This research also highlights the importance of *both* social and environmental uncertainty in driving agricultural outcomes. Finally, this research increases the capacity of water managers to monitor and predict changes in agricultural health. We constructed a user-friendly set of scripts that download, process and predict high resolution values of vegetation health for any MODIS tile. The tools we have constructed can be applied to any region in which MODIS data is collected. While this tool is best suited for regions with low cloud cover, it still performs well in one of the cloudiest regions of the world. The tool can be used to monitor and predict vegetation health at a high resolution in regions in which no local data is available, where it could support agricultural decision-making.

APPENDIX A: Objective functions

The farmers' objective functions follows Podesta et al. (2008). Under risk-averse utility farmers seek to maximize profit, but the utility of money declines the more one has (\$1000 would make a bigger difference to a person in poverty than to a millionaire). The utility of wealth or income w is given by:

$$u_{riskaverse}(w) = \begin{cases} \frac{w^{1-r} - 1}{1-r} & r \neq 1 \\ \ln(w) & r = 1 \end{cases}$$

where r is a coefficient of risk aversion. Larger values of r correspond to greater risk aversion, negative values to risk-seeking, and $r = 0$ represents indifference toward risk. Following the empirical literature, we use a value of $r = 1$, which corresponds to small risk aversion, with a constant gain in utility for each doubling of income or wealth (Podesta et al 2008).

Under regret-adjusted utility, farmers compare their profit to what might have happened had they made a different choice and the utility accounts for anticipated regret. This utility function determines the expected value of any one possible outcome by comparing it to all other possible outcomes. If the set of possible monetary outcomes (wealth or income) is $\{w_i\}$, then for a given outcome w^* , we define regret as the difference in risk-averse utility between w^* and the best possible outcome in $\{w_i\}$:

$$regret(w^*) = \max(u_{riskaverse}(w_i)) - u_{riskaverse}(w^*)$$

The regret-adjusted utility is given by:

$$u_{regretadjusted}(w) = u_{riskaverse}(w) - k(1 - \beta^{regret(w)})$$

where k sets the scale for regret and β ($0 \leq \beta < 1$) describes the decisionmaker's sensitivity to the magnitude of regret. Following the empirical literature, we set $r = 1$, $k = 0.155$, and $\beta = 0.5$ (Podesta et al. 2008).

Finally, under prospect theory, the utility of a given income does not depend on its magnitude, but on how much it exceeds or falls short of some reference value w_{ref} , with the pain of losses exceeding the pleasure in equal gains (Kahneman and Tversky 1979). When the consequence of a choice is uncertain, as described above, the prospect-theory utility is given by:

$$u_{prospect} = \sum f(\Delta w_i)g(p_i)$$

where

$$f(\Delta w) = \begin{cases} (\Delta w)^\alpha & \Delta w \geq 0 \\ -\lambda(-\Delta w)^\alpha & \Delta w < 0 \end{cases}$$

$$g(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$$

$$\gamma = \begin{cases} \gamma + & \Delta w \geq 0 \\ \gamma - & \Delta w < 0 \end{cases}$$

$\Delta w_i = w_i - w_{ref}$ is the change from the reference point (i.e. a farmer's past seasonal income), λ is the coefficient of loss aversion, α describes risk aversion/seeking and γ captures nonlinear probability weighting.

APPENDIX B: rSTAN code

```
ofc_model <- 'data {  
  
  int<lower=0> ngn;           //number of communities  
  int<lower=0> n;           //number of farmers  
  int<lower=0, upper=1> y[n];  
  int gn[n];               //mapping gn to group  
  vector[n] aw;  
  vector[n] major;  
  vector[n] female;  
  vector[n] sinhala;  
  vector[n] ses;  
  vector[n] land_owner;  
  vector[n] he;  
  vector[n] fo;  
}  
  
parameters{  
  real beta_aw;  
  real beta_major;  
  real beta_female;  
  real beta_sinhala;  
  real beta_ses;  
  real beta_land_owner;  
  real beta_he;  
  real beta_fo;  
  
  real alpha;  
  real<lower=0> sigma;  
  real a[ngn];             //group specific random effects  
}
```

```

model {

  beta_aw ~ cauchy(0,2.5);
  beta_major ~ cauchy(0,2.5);
  beta_female ~ cauchy(0,2.5);
  beta_sinhala ~ cauchy(0,2.5);
  beta_ses ~ cauchy(0,2.5);
  beta_land_owner ~ cauchy(0,2.5);
  beta_he ~ cauchy(0,2.5);
  beta_fo ~ cauchy(0,2.5);

  alpha ~ cauchy(20,20);
  sigma ~ cauchy(0, 10);
  a ~ normal(0, sigma);

  for (i in 1:n) {
    y[i] ~ bernoulli(inv_logit(alpha + a[gn[i]] + beta_aw*aw[i]
+ beta_major*major[i] +
                                beta_female*female[i] +
beta_sinhala*sinhala[i] + beta_ses*ses[i] +
beta_land_owner*land_owner[i] +
                                beta_he*he[i] +
beta_fo*fo[i]));
  }
}'

```


REFERENCES

- Anyamba, A. & Eastman, J.R., 1996. Interannual variability of NDVI over Africa and its relation to El Nino/Southern Oscillation. *International Journal of Remote Sensing*, 17(13), pp.2533–2548.
- Asoka, A. & Vimal, M., 2015. Prediction of vegetation anomalies to improve food security and water management in India. *Geophysical Research Letters*, 42(13), pp.5290–5298.
- Azmi, F., 2007. Changing livelihoods among the second and third generations of settlers in System H of the Accelerated Mahaweli Development Project (AMDP) in Sri Lanka. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, 61(1), pp.1–12. Available at: <http://www.tandfonline.com/doi/abs/10.1080/00291950601173903> [Accessed January 3, 2013].
- Bergstra, J.S., Yamins, D. & Davis, C., 2013. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. *Proceedings of the 30th International Conference on Machine Learning*, pp.115–123.
- Bolton, D.K. & Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology*, 173, pp.74–84.
- Brown, J.F. et al., 2002. A prototype drought monitoring system integrating climate and satellite data. *PECORA 15/Land Satellite Information IV/ISPRS Commission I/FIEOS 2002 Conference proceedings*. Available at: <http://www.isprs.org/proceedings/XXXIV/part1/paper/00074.pdf>.
- Burchfield, E.K. & Gilligan, J.G., 2016. Dynamics of individual and collective agricultural adaptation to water scarcity. *Winter Simulation Conference 2016*.
- Cai, X.L. & Sharma, B.R., 2010. Integrating remote sensing, census and weather data for an assessment of rice yield, water consumption and water productivity in the Indo-Gangetic river basin. *Agricultural Water Management*, 97(2), pp.309–316. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S037837740900290X> [Accessed March 23, 2015].
- Cash, D.W. et al., 2006. Scale and cross-scale dynamics: Governance and information in a multilevel World. *Ecology and Society*, 11(2), pp.8–20.
- Chandrasiri, J.K. & Bamunuarachchi, B.A., 2015. *Reasons for low adoption of selected OFC and vegetable varieties released by the Department of Agriculture*,
- Chattopadhyay, S. & Chattopadhyay, G., 2008. Comparative study among different neural net learning algorithms applied to rainfall time series. *Meteorological Applications*, 15(2),

pp.273–280.

- Chen, J. et al., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, 91(3–4), pp.332–344. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S003442570400080X> [Accessed January 23, 2014].
- Dai, A., 2013. Increasing drought under global warming in observations and models. *Nature Climate Change*, 3, pp.52–58.
- Dessai, S., 2009. Climate prediction: A limit to adaptation. In *Adapting to climate change: Thresholds, values, and governance*. pp. 64–78.
- Doraiswamy, P.C., Sinclair, T.R., Hollinger, S., Akhmedov, B., Stern, A., Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sensing of Environment*, 97, pp.192–202.
- Eastman, R., 1993. Evaluation Time Series Long Sequence Components Principal Standardized Using. *Photogrammetric Engineering & Remote Sensing*, 59(6), pp.991–996.
- Ellis, F., 1998. Household strategies and rural livelihood diversification. *The Journal of Development Studies*, 35(1), pp.1–38.
- Engle, N.L., 2011. Adaptive capacity and its assessment. *Global Environmental Change*, 21(2), pp.647–656. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0959378011000203> [Accessed July 10, 2014].
- Engle, N.L. & Lemos, M.C., 2010. Unpacking governance: Building adaptive capacity to climate change of river basins in Brazil. *Global Environmental Change*, 20(1), pp.4–13. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0959378009000466> [Accessed December 17, 2014].
- EPA, Region 9 Strategic Plan, 2011-14. 2014. Available at: <http://www3.epa.gov/region09/strategicplan/sanjoaquin.html>.
- Fluss, R., Faraggi, D. & Reiser, B., 2005. Estimation of the Youden Index and its associated cutoff point. *Biometrical Journal*, 47(4), pp.458–472.
- Galford, G.L. et al., 2008. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment*, 112(2), pp.576–587. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0034425707002258>.
- Gelman, A., 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in*

- Medicine*, 27, pp.2865–2873.
- Gelman, A. & Hill, J., 2007. *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press.
- Gibson, C.C., Ostrom, E. & Ahn, T.K., 2000. The concept of scale and the human dimensions of global change: a survey. *Ecological Economics*, 32(2), pp.217–239. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0921800999000920>.
- Giddens, A., 1984. *The constitution of society: Outline of the theory of structuration*, University of California Press.
- Government of Sri Lanka, 2010. *National climate change adaptation strategy for Sri Lanka - 2011 to 2016*, Colombo, Sri Lanka.
- Grimm, V.U. et al., 2010. The ODD protocol: A review and first update. *Ecological Modeling*, 221, pp.2760–2768.
- Gumma, M.K., 2011. Mapping rice areas of South Asia using MODIS multitemporal data. *Journal of Applied Remote Sensing*, 5(1), p.53547. Available at: <http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3619838> [Accessed February 4, 2014].
- Gunderson, L., 2001. *Panarchy: Understanding transformations in human and natural systems*, Island Press.
- Gupta, J. et al., 2010. The Adaptive Capacity Wheel: a method to assess the inherent characteristics of institutions to enable the adaptive capacity of society. *Environmental Science & Policy*, 13(6), pp.459–471. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1462901110000638>.
- Hanley, J.A. & McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), pp.29–36.
- Hansen, J.W., Marx, S.W. & Weber, E.U., 2004. *The role of climate perceptions, expectations, and forecasts in farmer decision making: The Argentine Pampas and South Florida: Final report of an IRI seed grant project*,
- Hastie, T., Tibshirani, R. & Friedman, J., 2009. *The elements of statistical learning: Data mining, inference and prediction*, Springer.
- Heim, R.R., 2002. A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83(8), pp.1149–1165.

- Holling, C.S., 2001. Understanding the complexity of economic, ecological, and social systems. *Ecosystems*, 4(5), pp.390–405.
- Holling, C.S. & Meffe, G.K., 1996. Command and control and the pathology of natural resource management. *Conservation Biology*, 10(2), pp.328–337. Available at: <http://onlinelibrary.wiley.com/doi/10.1046/j.1523-1739.1996.10020328.x/full> [Accessed December 10, 2014].
- Huete, A. et al., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), pp.195–213. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0034425702000962>.
- Imbulana, K.A.U.S., Wijesekera, N.T.S. & Neupane, B.R., 2006. *Sri Lanka National Water Development Report*,
- Jayawardene, H.K.W.I., Sonnadara, D.U.J. & Jayewardene, D.R., 2005. Trends of Rainfall in Sri Lanka over the Last Century. *Sri Lankan Journal of Physics*, 6, pp.7–17.
- Ji, L., Peters, A.J., 2004. Forecasting vegetation greenness with satellite and climate data. *IEEE Geoscience and Remote Sensing Letters*, 1(1), pp.3–8.
- Jinapala, K. et al., 2010. Volume 3: Policies , Institutions and Data Needs for Water Management. In *National Conference on Water, Food Security and Climate Change in Sri Lanka*. Colombo, Sri Lanka: International Water Management Institute.
- de Jong, I.H., 1989. *Fair and unfair: A study into the bethma system in two Sri Lankan village irrigation systems*, Colombo, Sri Lanka.
- Jönsson, P. & Eklundh, L., 2004. TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8), pp.833–845. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0098300404000974> [Accessed January 23, 2014].
- Kahneman, D. & Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(263–292).
- Kikuchi, M. et al., 2001. *Agro-well and pump diffusion in the dry zone of Sri Lanka*, Colombo, Sri Lanka.
- Kogan, F., Salazar, L. & Roytman, L., 2012. Forecasting crop production using satellite-based vegetation health indices in Kansas, USA. *International Journal of Remote Sensing*, 33(9), pp.2798–2814.

- Koide, N. et al., 2012. Prediction of rice production in the Philippines using seasonal climate forecasts. *Journal of Applied Meteorology and Climatology*, 52(3), pp.552–569.
- Kolstad, C.D., 2011. *Environmental economics* 2nd Ed., Oxford University Press.
- Lasaponara, R., 2006. On the use of principal component analysis (PCA) for evaluating interannual vegetation anomalies from SPOT/VEGETATION NDVI temporal series. *Ecological Modelling*, 194(4), pp.429–434. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0304380005005454> [Accessed November 10, 2014].
- Lebel, L., Garden, P. & Imamura, M., 2005. The politics of scale, position, and place in the governance in the Mekong region. *Ecology and Society*, 10(2).
- Lesk, C., Rowhani, P. & Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. *Nature*, 529, pp.84–87.
- Lin, B., 2011. Resilience in agriculture through crop diversification: Adaptive management for environmental change. *BioScience*, 61(3), pp.183–193.
- Liu, J. et al., 2007. Complexity of coupled human and natural systems. *Science (New York, N.Y.)*, 317(5844), pp.1513–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17872436> [Accessed March 5, 2013].
- Ltd., S.I.P., 2013. *Updated Mahaweli Water Resources Development Plan*, Colombo, Sri Lanka.
- Lupo, F.M. et al., 2007. Categorization of land-cover change processes based on phenological indicators extracted from time series of vegetation index data. *International Journal of Remote Sensing*, 28(11), pp.2469–2483.
- Mahaweli Authority of Sri Lanka, 2012. *Statistical Handbook*, Colombo, Sri Lanka.
- Malmgren, B. a. et al., 2003. Precipitation trends in Sri Lanka since the 1870s and relationships to El Nino-southern oscillation. *International Journal of Climatology*, 23(10), pp.1235–1252. Available at: <http://doi.wiley.com/10.1002/joc.921> [Accessed January 15, 2015].
- Manthrithilake, H. & Liyanagama, B., 2012. Simulation model for participatory decision making: water allocation policy implementation in Sri Lanka. *Water International*, (September), pp.37–41. Available at: <http://www.tandfonline.com/doi/abs/10.1080/02508060.2012.708602> [Accessed July 3, 2013].

- Manthritilake, H. & Liyanagama, B.S., 2012. Simulation model for participatory decision making: water allocation policy implementation in Sri Lanka. *Water International*, 37(4), pp.478–491.
- MASL, 2014. *Yala 2014 Seasonal Summary Report*,
- Meinzen-Dick, R., 2007. Beyond panaceas in water institutions. *Proceedings of the National Academy of Sciences of the United States of America*, 104(39), pp.15200–5. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2000530&tool=pmcentrez&rendertype=abstract> [Accessed March 19, 2015].
- Mo, X. et al., 2005. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. *Ecological Modelling*, 183, pp.301–322.
- Moss, T. & Newig, J., 2010. Multilevel water governance and problems of scale: setting the stage for a broader debate. *Environmental management*, 46(1), pp.1–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20640851> [Accessed January 15, 2015].
- Nations, U., 2006. *World population prospects: The 2004 Revision*, New York.
- Nordhaus, W.D., 2008. *A question of balance: Weighing the options on global warming policies.*, Yale University Press.
- Ostrom, E., 2009. A general framework for analyzing sustainability of social-ecological systems. *Science (New York, N.Y.)*, 325(5939), pp.419–22. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19628857> [Accessed October 26, 2012].
- Pahl-wostl, C. et al., 2007. Social Learning and Water Resources Management. , 12(2).
- Peters, A.J. et al., 2002. Drought Monitoring with NDVI-Based Standardized Vegetation Index. *Photogrammetric Engineering & Remote Sensing*, 68(1), pp.71–75.
- Podesta, G.E. et al., 2008. Agricultural decision making in the Argentine Pampas: Modeling the interaction between uncertain and complex environments and heterogeneous and complex decision makers. In T. Kugler et al., eds. *Decision Modeling and Behavior in Complex and Uncertain Environments*. Springer, pp. 57–76.
- Prasanna, R.P.I.R., Bulakulama, S.W.G.K. & Kuruppuge, R.H., 2011. Factors affecting farmers' higher grain from paddy marketing: A case study on paddy farmers in North Central Province, Sri Lanka. *International Journal of Agricultural Management and Development*, 2(1), pp.57–69.

- Ranasinghe, D.M.S.H.K., 2013. Environmental consequences of Moragahaka, NDA development project. In *Proceedings of International Forestry and Environment Symposium*. Available at: <http://journals.sjp.ac.lk/index.php/fesympo/article/view/1643>.
- Rasmussem, M.S., 1992. Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. *International Journal of Remote Sensing*, 13(18), pp.3431–3442.
- Rayner, N.A. et al., 2003. Global analysis of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108(D14), p.4407.
- Rhee, J., Im, J. & Carbone, G.J., 2010. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sensing of Environment*, 114(12), pp.2875–2887. Available at: <http://dx.doi.org/10.1016/j.rse.2010.07.005>.
- Richards, J.A., 2013. *Remote sensing digital image analysis: An introduction* 5th ed., New York: Springer.
- Sakamoto, T. et al., 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, 96(3–4), pp.366–374. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0034425705001057> [Accessed January 27, 2014].
- Samad, M., 2005. Water institutional reforms in Sri Lanka. *Water Policy*, 7, pp.125–140.
- Savitzky, A. & Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified lead squares procedures. *Analytical Chemistry*, 36(8), pp.1627–1693.
- Senaratne, A. & Scarborough, H., 2011. Coping with climate variability by rain-fed farmers in Dry Zone, Sri Lanka: Towards understanding adaptation to climate change. In *AARES: Australian Agricultural & Resource Economics Society 55th Annual Conference Handbook*. Melbourne, Australia, pp. 1–22.
- Shah, T. et al., 2013. Ancient small-tank irrigation in Sri Lanka. *Economic & Political Weekly*, xlviII(11), pp.58–63. Available at: http://www.epw.in/system/files/pdf/2013_48/11/Ancient_SmallTank_Irrigation_in_Sri_Lanka.pdf [Accessed April 30, 2013].
- Shah, T. et al., 2003. Sustaining Asia’s groundwater boom: An overview of issues and evidence. *Natural Resources Forum*, 27(2), pp.130–141. Available at: <http://doi.wiley.com/10.1111/1477-8947.00048>.

- De Silva, C.S. et al., 2007. Predicting the impacts of climate change—A case study of paddy irrigation water requirements in Sri Lanka. *Agricultural Water Management*, 93(1–2), pp.19–29. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S037837740700162X>.
- Singh, P. & Borah, B., 2013. Indian summer monsoon rainfall prediction using artificial neural network. , pp.1585–1599.
- Small, C., 2012. Spatiotemporal dimensionality and Time-Space characterization of multitemporal imagery. *Remote Sensing of Environment*, 124, pp.793–809. Available at: <http://dx.doi.org/10.1016/j.rse.2012.05.031>.
- Small, C. & Milesi, C., 2013. Multi-scale standardized spectral mixture models. *Remote Sensing of Environment*, 136, pp.442–454. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0034425713001788> [Accessed January 28, 2015].
- Smit, B. & Wandel, J., 2006. Adaptation, adaptive capacity and vulnerability. *Global Environmental Change*, 16(3), pp.282–292. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0959378006000410> [Accessed January 28, 2013].
- Spiertz, H.L.J. & de Jong, I.J.H., 1992. Traditional law and irrigation management: The case of bethma. In G. Diemer & J. Slabbers, eds. *Irrigators and engineers: Essays in honour of Lucas Horst*. Amsterdam, NL: Thesis Publishers, pp. 185–201.
- Swain, D.L. et al., 2014. The extraordinary California drought of 2013/2014: Character, context, and the role of climate change. *Bulletin of the American Meteorological Society*, 95(7), pp.S3–S7.
- Takesada, N., Manatunge, J. & Herath, I.L., 2008. Resettler choices and long-term consequences of involuntary resettlement caused by construction of Kotmale Dam in Sri Lanka. *Lakes & Reservoirs: Research & Management*, 13(3), pp.245–254. Available at: <http://doi.wiley.com/10.1111/j.1440-1770.2008.00374.x> [Accessed November 21, 2012].
- Thenkabail, P., Gamage, M. & Smakhtin, V., 2004. *The use of remote sensing data for drought assessment and monitoring in Southwest Asia*, Colombo, Sri Lanka. Available at: http://books.google.com/books?hl=en&lr=&id=BiG6G4am-WEC&oi=fnd&pg=PR5&dq=The+Use+of+Remote+Sensing+Data+for+Drought+Assessment+and+Monitoring+in+Southwest+Asia&ots=FJMokpUD2N&sig=Sax4TznnZCIRJdZ_N9ICu6SaHQw [Accessed October 27, 2013].
- Thiruchelvam, S., 2010a. Agricultural production efficiency of bethma cultivation in Mahaweli System H. *Sri Lankan Journal of Agricultural Economics*, 7. Available at: <http://www.sljol.info/index.php/SJAE/article/view/1820>.

- Thiruchelvam, S., 2010b. Agricultural production efficiency of bethma cultivation in Mahaweli System H. *Sri Lankan Journal of Agricultural Economics*, 7, pp.1–20. Available at: <http://www.sljol.info/index.php/SJAE/article/view/1820>.
- Tversky, A. & Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, pp.297–323.
- Villholth, K.G. & Rajasooriyar, L.D., 2009. Groundwater resources and management challenges in Sri Lanka - an overview. *Water Resources Management*, 24(8), pp.1489–1513. Available at: <http://www.springerlink.com/index/10.1007/s11269-009-9510-6> [Accessed March 13, 2013].
- WFP, 2014. *Rapid food security assessment in districts affected by erratic weather conditions in Sri Lanka*,
- Wilhite, D. & Vanyarkho, D., 2000. Drought: Pervasive impacts of a creeping phenomenon. In D. Wilhite, ed. *Drought: A global assessment*. London: Routeledge, pp. 245–255.
- Withananachchi, S.S. et al., 2014. Water resource management in dry zonal paddy cultivation in Mahaweli River Basin, Sri Lanka: An analysis of spatial and temporal climate change impacts and traditional knowledge. *Climate*, 2(4), pp.329–354.
- Xiao, X. et al., 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, 100(1), pp.95–113. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0034425705003433> [Accessed August 8, 2011].
- Ziervogel, G. et al., 2010. Using climate information for supporting climate change adaptation in water resource management in South Africa. *Climatic Change*, 103(3–4), pp.537–554.
- Zubair, L., 2005. Modernisation of Sri Lanka's Traditional Irrigation Systems and Sustainability. *Science Technology & Society*, 10(2), pp.161–195.