ESTIMATING SOURCES OF VALLEY FEVER PATHOGEN PROPAGATION IN SOUTHERN ARIZONA: A REMOTE SENSING APPROACH

By

Frederick S. Pianalto

_____

A Dissertation Submitted to the Faculty of the

SCHOOL OF GEOGRAPHY AND DEVELOPMENT

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2013

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Frederick S. Pianalto, entitled *Estimating Sources of Valley Fever Pathogen Propagation in Southern Arizona: A Remote Sensing Approach*, and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: (11/26/2013)
Andrew C. Comrie

_____ Date: (11/26/2013)
Charles F. Hutchinson

_____ Date: (11/26/2013)
Stuart E. Marsh

_____ Date: (11/26/2013)
Stephen R. Yool

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 12/04/2013
Dissertation Director: Stephen R. Yool

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Frederick S. Pianalto

ACKNOWLEDGMENTS

The incredible Geography and Development staff, including Linda, John, Gabe, Maria, Tawny, and Liz, and Alicia at the GIDP, have reliably provided me with abundant support over the years, in terms of time, equipment, funding, and cheer. Hats off to them! My dissertation committee, Dr. Comrie, Dr. Hutchinson, Dr. Marsh, and Dr. Yool, ensured that my comprehensive exams and dissertation defense were challenging, and rewarding experiences. I am genuinely impressed with their attention and helpful critique. I extend many thanks to Dr. Yool. As my advisor, he has been a patient sounding board for my research travails, and he provided just the right amount of advice. I will be at loss without our weekly progress meetings!

TABLE OF CONTENTS

ABSTRACT

Coccidioidomycosis (Valley Fever) is an environmentally-mediated respiratory disease caused by the inhalation of airborne spores from the fungi *Coccidioides spp.* The fungi reside in arid and semi-arid soils of the Americas. The disease has increased epidemically in Arizona and other areas within the last two decades. Despite this increase, the ecology of the fungi remains obscure, and environmental antecedents of the disease are largely unstudied. Two sources of soil disturbance, hypothesized to affect soil ecology and initiate spore dissemination, are investigated.

Nocturnal desert rodents interact substantially with the soil substrate. Rodents are hypothesized to act as a reservoir of coccidioidomycosis, a mediator of soil properties, and a disseminator of fungal spores. Rodent distributions are poorly mapped for the study area. We build automated multi-linear regression models and decision tree models for ten rodent species using rodent trapping data from the Organ Pipe Cactus National Monument (ORPI) in southwest Arizona with a combination of surface temperature, a vegetation index and its texture, and a suite of topographic rasters. Surface temperature, derived from Landsat TM thermal images, is the most widely selected predictive variable in both automated methods.

Construction-related soil disturbance (e.g. road construction, trenching, land stripping, and earthmoving) is a significant source of fugitive dust, which decreases air quality and may carry soil pathogens. Annual differencing of Landsat Thematic Mapper (TM) mid-infrared images is used to create change images, and thresholded change areas are associated with coordinates of local dust inspections. The output metric identifies

source areas of soil disturbance, and it estimates the annual amount of dust-producing surface area for eastern Pima County spanning 1994 through 2009.

Spatially explicit construction-related soil disturbance and rodent abundance data are compared with coccidioidomycosis incidence data using rank order correlation and regression methods. Construction-related soil disturbance correlates strongly with annual county-wide incidence. It also correlates with Tucson periphery incidence aggregated to zip codes. Abundance values for the desert pocket mouse (*Chaetodipus penicillatus*), derived from a soil-adjusted vegetation index, aspect (northing) and thermal radiance, correlate with total study period incidence aggregated to zip code.

CHAPTER 1: INTRODUCTION

1.1 Introduction

1.1.1 Disciplines

Health geography is the application of spatial information, perspectives, and methods to the study of health, disease and health care. Health is defined as "a state of complete physical, mental, and social well-being and not merely an absence of disease" (WHO, 2006). Health geography, a broad discipline, it is situated in a complex interplay of environmental factors, including physical, biological, economic, social and cultural dimensions. Historically, health geography is considered to be divided into two streams: 1) the geographies of health and disease, with descriptive and analytical research quantifying disease frequencies, distributions, and characteristics of susceptibility; and 2) the geography of health care, such as facility location, accessibility and utilization (Meade and Earickson, 2000).

Medical geography, as a subdiscipline, is the study of the relation between geographic factors and disease (Philo, 2009). Inherent is an after-the-fact positioning for treatment and the study of disease. Medical geography focuses on "who has what disease where" (Meade and Earickson, 2000). Historically, it has been treated as a tool for the application of geographical methods to medical problems. Litva and Eyles (1995) assert that its atheoretical nature is merely an appearance; it has been influenced by positivist philosophy, where knowledge and information is derived from logical, mathematical treatments and empirical evidence. Medical geography has become increasingly concerned with health geography as a behavioral and social construction, and disease

ecology as an interface between the natural, physical world and the cultural dimension of existence.

Epidemiology, a science or discipline in its own right, focuses on the prevalence and spread of disease in a community and on the health of human populations. It is often considered the key scientific underpinning of public health. It generally is concerned with strict study design and the set-up of case control and cohort frameworks. Spatial epidemiology, as a branch of epidemiology, is concerned with describing, quantifying, and explaining geographical variations in disease, especially with respect to variations in environmental exposures at the small-area scale (Elliot et al. 2000). Examples of their foci are disease mapping, comparisons of risk maps with exposure maps, geographical correlation studies, disease cluster detection, spatial surveillance, and provision of information on the health needs of a population.

1.1.2 Technology

A geographic information system (GIS), with its ability to manage and portray spatial data, has become the dominant tool in geography. A geographic information system (GIS) is an integrated set of tools and methodologies for collecting, storing, retrieving, analyzing, and displaying spatial as well as non-spatial attribute data. It has transformed a variety of health analyses and the structuring of public data (Meade and Earickson, 2000).

Perhaps the most important source for the development of the GIS, and its data, is remote sensing (Meade and Earickson, 2000). Satellites provide an enormous amount of digital data in multiple bands of the electromagnetic spectrum. Continuous and

repetitious image collection allows differentiation of land cover or usage, and provides a platform for monitoring for change. Methodologies for the processing of remotely-sensed images are a source of a variety of raster data products.

1.1.3 Coccidioidomycosis

Coccidioidomycosis (Valley Fever) is a systemic infection characterized by fever, respiratory infection and reddish bumps on the skin. It is caused by inhalation of airborne spores from *Coccidioides immitis* and *Coccidioides posadasii* fungi, which are endemic in the southwestern United States, and in parts of Mexico, Central and South America. Coccidioidomycosis is produced in Arizona by inhalation of spores from *C. posadasii* (Fisher et al., 2002).

The *Coccidioides spp.* reside in warm, arid and semi-arid soils of the Americas. The fungi exist in a dimorphic life cycle consisting of saprophytic mycelial phase in the soil, and a parasitic spherule phase when arthroconidia are inhaled by a mammalian host (Cole and Sun, 1985). Given proper conditions, slender filaments of saprophytic cells (hyphae) grow in the upper part of the soil (Kolivras, 2001). After a week or more of growth, many of the hyphal cells mature into rectangular arthrospores. The arthrospores alternate with smaller, sterile cells. When the soil dries, the alternating sterile cells breach easily, freeing the intervening arthrospores. The arthrospores range from 1.5 to 4.5 μm in width and 5.0 to 30 μm in length.

The parasitic stage usually initiates in the lungs, and it can spread to other parts of the body. If inhaled, the arthrospores can penetrate to the smallest bronchiole or pulmonary alveoli of the lung. In the parasitic phase, the arthrospore or arthroconidia

develops into a spherical, double-walled cell called a spherule (sporangium). Spherules measure from 10 to 200 µm, and they typically contain a few to several hundred endospores, each 2 to 5 µm in diameter. The spherule eventually ruptures, discharging the endospores into the neighboring tissue. Each endospore is potentially capable of blooming into a new spherule.

Exposure usually occurs following events that disrupt the soil, resulting in aerosolization of the fungal arthrospores (Schneider et. al., 1997). Both natural and anthropogenic or human soil disturbance can produce spore dispersion. Hypothesized sources of soil disturbance include the wind, dust devils, construction-related activities, agriculture, archeological digs, and rodent activity.

1.1.4 Study Area

The study is situated in the Sonoran Desert, an arid region stretching from southeastern Arizona, across Sonora, Mexico and through most of Baja California. The Sonoran Desert is characteristic for its semi-arid climate, mild winters and a bimodal rainfall pattern. Mountain ranges, volcanic hills, bajadas (coalesced alluvial fans), valley floors, washes and arroyos (steep-sided gulches) are typical of its geomorphology (Hoffmeister, 1986). The visually dominant elements of the landscape are two characteristic life forms: legume trees and large columnar cacti (Arizona-Sonora Desert Museum, 2013).

The Lower Sonoran Life Zone (LSLZ), based on Merriam's elevation-precipitation life zones for the southwest (Merriam and Steineger, 1890), further delineates the study area. The LSLZ are the lower elevation parts of the desert (generally

< 1000 m). The LSLZ stretches, in Pima County, Arizona, from the relatively pristine, sparsely-populated Organ Pipe Cactus National Monument, situated in the western half of the county, to the Tucson metropolitan area in the eastern half.

A ring of communities, unincorporated urban development, and undeveloped areas surround the city of Tucson. Here, thousands of acres of the Sonoran Desert are at the forefront of urbanization, and hundreds of acres have been bladed for the construction of houses and commercial strip malls (AIA, 2007). Between 1980 and 1990, the city's area increased by over 50 percent to approximately 600 km^2 through the annexation of unincorporated land. The regional population has also experienced a doubling in a similar time period. The population of Pima County is approximately one million (Census, 2012).

1.2 Explanation of the Problem

1.2.1 Impacts of coccidioidomycosis

Most cases of *Coccidioides spp.* infections are self-resolved. Approximately 40% of infections are symptomatic, and they can result in severe complications such as influenza-like illnesses, community acquired pneumonia (CAP), lung cavities, and disseminated infections in the central nervous system, skin, bones, joints and other organs (Komatsu et al., 2003). About one percent of infected individuals experience serious, life threatening conditions such as meningitis and organ damage. On average, patients diagnosed with the disease suffer symptoms for six months. Risk factors for severe infections include race, age, and immunosuppression. The fungi also infect livestock, pets and wild animals.

Rates of coccidioidomycosis have increased epidemically in Arizona and other areas within the last two decades. The Arizona Department of Health Services (ADHS) reported a coccidioidomycosis incidence of 43 cases per 100,000 of population in 2001, representing an increase of 186% since 1995 (Komatsu et al. 2003). In 1998, 1,551 cases were reported to the ADHS; 5,535 were reported in 2006, and over 10,000 were reported in 2009 (Komatsu et al. 2003; Sunenshine et al. 2007; Hector et al. 2011). It is the fourth most common disease reported to the Arizona Department of Health Services (Park et al. 2005), and Arizona accounts for 60 percent of reported cases in the nation (Center for Diseases Control and Prevention, 2004). Most (95%) of Arizona cases are in Maricopa, Pinal, and Pima Counties.

Coccidioidomycosis became a nationally reportable disease in 1995 at the southwest regional level, at which time a case definition was adopted that required laboratory confirmation. The reporting requirement is suggested to play a role, at least partially, in the increasing linear trend in exposure rates over the previous two decades (Sunenshine et al. 2007, Tamerius and Comrie, 2011).

The consequent public health burden to affected regions is considerable. Hospital charges for coccidioidomycosis in the U.S. exceeded $86 million in 2007. Recent studies noted that direct hospital charges for the treatment of the disease in Arizona was 26.8 million dollars with a median charge of $14,292 (Nguyen et al., 2013). It is an increasingly important health issue due to migration into the state from other regions and increased numbers of immuno-suppressed patients.

1.2.2 Investigations of coccidioidomycosis sources

Environmental factors are hypothesized to explain the occurrence and distribution of *Coccidioides spp.* at all points in the disease chain. Despite the significant increase in incidence of coccidioidomycosis, and the associated interest due to its economic and human costs, there is limited knowledge about the specific ecological niche required for *Coccidioides spp.* to flourish, the environmental antecedents of disease outbreaks, and the precise mechanisms of spore aerosolization and subsequent dispersion (Galgiani, 1999).

Coccidioidomycosis cannot be transmitted from person to person, but it is acquired by inhalation of the arthrospores from the environment. Therefore it is reasonable to assume that the fungus should be easily isolated from endemic areas (Ajello et al. 1965; Lacy & Swatek 1974). However, only a few positive isolations from environmental samplings have been obtained in highly endemic areas in the United States (Stewart and Meyer 1932; Emmons 1942; Maddy 1965). The scarce environmental evidence for *Coccidioides spp.* seems to be in disagreement with the high incidence rates obtained for the disease. Detection of the fungus in the environment remains a critical challenge to modeling the source of disease (Barker et al., 2012).

On a regional scale, the major predictors of disease are climate, soil disturbance, and dust or wind events (Pappagianis 1994; Comrie 2005; Comrie and Glueck 2007, Tamerius and Comrie, 2011). Stacy et al. (2012) used a remote sensing time series to relate moist soils in the early spring, resulting from antecedent winter precipitation, with increased coccidioidomycosis incidence up to a year later. However, at finer scales, the ecology of the fungus remains obscure and largely unstudied (Cox and Magee, 2004).

Infection usually occurs following activities or natural events that disrupt the soil, resulting in aerosolization of the fungal arthrospores (Schneider et. al., 1997). Wind erosion may contribute to spore dispersion. Evidence suggests that anthropogenic or human-induced soil disturbance generates dispersal of spores. Outbreaks of the disease have been associated with soil disruption, archeological digs, agriculture, and construction (Cairns et al., 2000; Park et al. 2005, Fisher et al., 2007). *Coccidioides spp.* is associated with alkaline soil that has a high salt content, rodents, rodent burrows and Amerindian middens (Swatek, 1970).

1.2.3 Public health challenges

Due to mandatory reporting, public health agencies in the U.S. have effectively established a surveillance system for coccidioidomycosis in endemic areas. Dissemination of information, another public health mandate, on the risks and control of coccidioidomycosis is improving.

However, agencies are not in position to provide detailed exposure and control information. Recommendations are limited to avoid blowing dust and activities that disturb soil, and to use dust masks and respirators in these situations. Evidenced by a lack of risk assessments, models, vulnerability maps and early warning systems, they have been unable to link coccidioidomycosis surveillance data with information on environmental factors explaining the occurrence and distribution of the disease.

Ultimately, improvement in the control of coccidioidomycosis incidence will require research into the molecular and cellular biology of *Coccidioides spp.*, vaccine development, discovery of new antifungal drugs, and a better understanding of the soil

ecology that supports the fungus in its endemic regions and the mechanisms of spore dispersal (Galgiani, 1999).

1. 3 Research Questions

The dissertation is organized around three primary research questions, summarized below. Through these, the relationship between two sources of soil disturbance, hypothesized to affect soil ecology and initiate *Coccidioides spp.* spore dissemination, and coccidioidomycosis incidence are investigated.

A. What is the spatial distribution of burrowing desert rodent abundance in Pima County, Arizona?

- What biophysical variables can be estimated by remote sensing, Geographic Information System (GIS), and statistical methods to estimate burrowing rodent abundance?
- How can these variables be incorporated into a predictive model?

B. What are the spatial and temporal characteristics of construction-related soil disturbance in Pima County, Arizona?

- Can construction-related disturbance be measured by remote sensing and GIS methods?
- Is this metric related to dust generation?

C. What is the spatial relationship between desert burrowing rodent abundance distributions and construction-related soil disturbance metrics with coccidioidomycosis (Valley Fever) incidence in Pima County, Arizona?

- What kind or type of modeling can support these data?
- How does the relationship between coccidioidomycosis incidence and a combined habitat-disturbance model vary across spatial scale or size of enumeration unit?

1.4 Approach

Spatial sciences are invoked to provide a new dimension in characterizing coccidioidomycosis and its contributing factors. Spatial tools available to the health geographer, such as Geographical Information Systems (GIS), remote sensing, epidemiology, and statistics, are employed to achieve these goals. The convergence of geography with public health, ecology and environmental sciences are utilized to achieve new perspectives.

A top-down research approach is employed. Available data sets and techniques are investigated to determine what research questions can be asked of the phenomena in terms of their scale, and research questions were repositioned to match availability and scale of data. For example, the resolution of the Landsat TM (30m pixel) is suited for the evaluation of rodent plots (90m) and construction disturbance (average 480m dimension). The extent of Landsat TM imagery (185 km scenes) is appropriate for data storage and processing requirements of a regional study. As another example, coccidioidomycosis incidence data are available at zip code granularity; this may be a sufficient match for a spore dissemination model from a disturbance site. Temporal comparison periods are

limited by repeat satellite visits and seasonal phenology. These scales define, enhance and constrain methodological capabilities.

Multiple data and methods are used to enhance the understanding of phenomena of study, triangulate research questions and validate output. For example, fugitive dust inspection data and rodent abundance data are obtained from public agencies, and information available in previous scientific literature is collected. These are compared with research output to drive model building and to validate results. Comparisons of parametric and non-parametric regression techniques with identical data input are used to instruct rodent abundance models. Variations in spatial and temporal scales, for example, county and zip code areas, and multi-year and annual periods, are used to ascertain relations of coccidioidomycosis incidence and soil disturbance.

The work presented here explores the use of viable, readily available, and free data sources generated by remote sensing methods and GIS methods. Satellite imagery is obtained from the Landsat 5 Thematic Mapper (TM) data archive. The Landsat 5 TM provides data since 1984 in seven spectral regions, including the mid-infrared (NASA, 2011). The data is processed to a standard output, and it is freely available from USGS data banks. Digital elevation models of the study area are also obtained at no cost from U.S. Geological Service data sources (USGS, 2013). Other ancillary data are researched and explored for assistance in modeling at all study phases. Political, census, agriculture, soil, and vegetation spatial data are obtained at no cost from online data sources.

Key data sets were obtained by information request from public agencies: county fugitive dust inspection data, national monument nocturnal rodent trapping data, and state coccidioidomycosis incidence data. Availability and access to these data were

instrumental for the study, and without them this research project would not have been accomplished. Likewise, methods and models developed here will be reported and made available as tools to public health agencies charged with characterizing and preventing the disease. To facilitate use by others, strategies are designed for minimal data processing for achievement of satisfactory results.

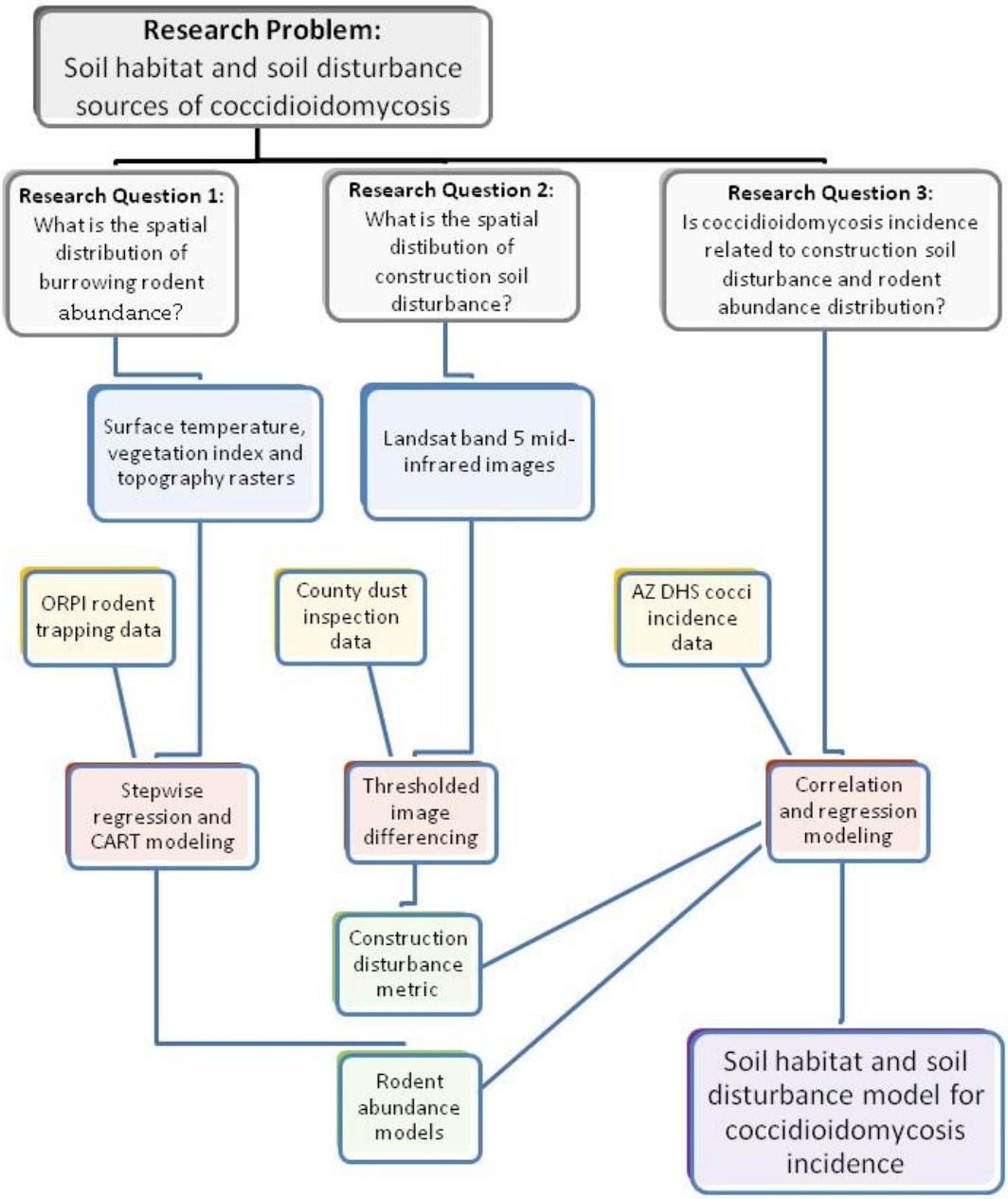
1.5 Explanation of the Dissertation Format

This dissertation offers new knowledge pertaining to the environmental sources of *Coccidioides spp.* dissemination. Three papers present important and interrelated insights into this problem (Figure 1).

Key components of this dissertation are formatted as scientific journal submissions. They are presented as Appendices A through C. The first article will be submitted to *Ecological Modelling* or a similar publication. The second article will be submitted to *GIScience & Remote Sensing*. The third article will be submitted to *Emerging Infectious Diseases* or a similar publication. Chapter 2 follows this introductory chapter with a brief summary of the three articles.

I am principally responsible for formulating research questions, conducting data analysis, and writing manuscripts. This dissertation represents my original and individual efforts, and I will be the lead author on the publications arising from this work.

Figure 1: Research plan



CHAPTER 2: PRESENT STUDY

The methods, results, and conclusions of the present study are detailed in three manuscripts appended to this dissertation. A brief summary of key features and findings follows.

2.1 Nocturnal desert rodent abundance distribution in southwestern Arizona estimated by regression methods with remote sensing-derived data.

Nocturnal rodents play a key role in the Sonoran Desert ecosystem as consumers and prey. They interact substantially with the soil substrate and alter soil properties. They are also linked to the transmission of human diseases. Rodent distributions are poorly mapped for the study area, and factors derived from remotely-sensed data that explain rodent biogeography are not known. We build automated multi-linear regression models and decision tree models for ten rodent species using rodent trapping data from the Organ Pipe Cactus National Monument (ORPI) in southwest Arizona with a combination of surface temperature, a vegetation index and its texture, and a suite of topographic rasters as explanatory variables. The parametric and non-parametric models show remarkable congruency between themselves and with outside sources. Surface temperature is the most widely selected explanatory variable in both automated methods; we attribute this to its integrating capabilities of multiple landscape characteristics. Model output is used to develop spatially explicit distribution maps.

2.2 Fugitive dust emission sources arising from construction: a remote sensing approach.

Construction-related soil disturbance (e.g., road construction, trenching, land stripping, earthmoving and blasting), is a significant source of fugitive (airborne) dust in the atmosphere. Fugitive dust is a primary cause of decreased air quality and may carry airborne pathogens. We use Landsat Thematic Mapper (TM) remote sensing data spanning 1994 through 2009 over southern Arizona to identify source areas of construction-related activity likely to produce fugitive dust. We correlate temporal changes in the mid-infrared spectral response to dust sources from local construction. Image differencing of the TM band 5 (mid-infrared), with a change threshold of ± five standard deviations of the mean, suitably estimates the location and area affected by construction-related soil disturbance. Estimated dust-producing surface area ranges from 10.0 km^2 (1996-1997) to 28.3 km^2 (2004-2005), or 0.16% to 0.44% of the Pima County study area. Our methods aim to automate monitoring of fugitive dust sources by environmental and health agencies and to provide inputs to dust transport, air quality and climate models.

2.3 Estimating environmental sources of Valley Fever propagation in southern Arizona with remote sensing.

Coccidioidomycosis (Valley Fever) is an environmentally-mediated respiratory disease caused by the inhalation of airborne spores from the fungi *Coccidioides spp*. The fungi reside in arid and semi-arid soils of the Americas. The disease has increased epidemically in Arizona and other areas within the last decade. Despite this increase, the

ecology of the fungus remains obscure, and environmental antecedents of the disease are largely unstudied. We investigate two sources of soil disturbance, anthropogenic construction and nocturnal desert rodents, which are hypothesized to affect soil ecology and initiate *Coccidioides spp.* spore dissemination. We estimate construction-related soil disturbance with annual differencing of Landsat Thematic Mapper mid-infrared images. Source areas of soil disturbance are identified, and annual affected areas are estimated for eastern Pima County, Arizona and for zip code areas spanning 1994 through 2006. We build rodent abundance distribution maps for the study area using regression models of biophysical variables derived from remotely-sensed data with comparisons to rodent trapping data from the Organ Pipe Cactus National Monument. The two spatially explicit soil disturbance sources are compared with coccidioidomycosis incidence data using rank order correlation and regression methods. Construction-related soil disturbance correlates with annual county-wide incidence (R^2 = 0.49, p-value 0.012), and with incidence of zip codes at the periphery of the city of Tucson for the total study period (R^2 = 0.48, p-value 0.001). The average abundance values for the desert pocket mouse (*Chaetodipus penicillatus*), derived from a soil-adjusted vegetation index, aspect (northing) and thermal radiance, correlate with total study period incidence aggregated to zip code (R^2 = 0.25, p-value 0.02).

2.4 Future work

Additional investigations might be undertaken to improve and extend the results of this study. Construction change pixels or data can be weighted by its corresponding rodent abundance to produce a combined model for incidence prediction. Satellite

imagery and methods can be performed to ascertain temporal variations in rodent abundance distributions, such as seasonal and inter-annual, for comparisons with incidence data. Data sets for the study ended at 2006; recent satellite and incidence data can be obtained to extend the lifespan of the study. Masks can be applied to limit the analysis to the periphery of Tucson, where there is a more solid theoretical basis for linking construction and rodents to coccidioidomycosis incidence. As a final example, the study and methods can be extended to Maricopa and Pinal counties in Arizona, which are experiencing similar urban growth and Valley Fever incidence dynamics.

REFERENCES

AIA, 2007, Tucson/Pima County, Arizona: One Million Reasons to Plan for Sustainable Growth, Southern Arizona Chapter of the American Institute of Architects, http://www.aia.org/aiaucmp/groups/aia/documents/pdf/aias078080.pdf.

Ajello, L., Maddy, K., Crecelius, G., Hugenholtz, P.G., and L.B. Hall, 1965. Recovery of Coccidioides immitis from the air, Journal of the International Society for Human and Animal Mycology, 4(2): 92-95.

Arizona-Sonora Desert Museum, 2000. A natural history of the Sonoran Desert, Phillips, S.J. and P.W. Comus, Eds., University of California Press, Berkeley CA, 650 pp.

Barker, B.M., Tabor, J.A., Shubitz, L.F., Perrill, R., and M.J. Orbach, 2012. Detection and phylogenetic analysis of Coccidioides posadasii in Arizona soil samples, Fungal Ecology 5: 163-176.

Cairns, L., Blythe, D., Kao, A., Pappagianis, D., Kaufman, L., Kobayashi, J. and R. Hajjeh, 2000. Outbreak of coccidioidomycosis in Washington state residents returning from Mexico, Clinical Infectious Diseases, 30: 61-64.

Census, 2012, State & County QuickFacts, Pima County, Arizona, U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/04/04019.html.

Cole, G.T. and S.H. Sun, 1985. Arthroconidium - spherule - endospore transformation in Coccidioides immitis. In: Fungal dimorphism, Szaniszlo P.J. (Ed), Plenum Publishing Corporation, New York City 1985. p.281.

Comrie, A.C., 2005. Climate factors influencing coccidioidomycosis seasonality and outbreaks. Environmental Health Perspectives, 113: 688–692.

Comrie A.C. and M.F. Glueck, 2007. Assessment of climate-coccidioidomycosis model: model sensitivity for assessing climatologic effects on the risk of acquiring coccidioidomycosis. Annals of the New York Academy of Sciences 1111: 83–95. doi: 10.1196/annals.1406.024.

Cox, R.A. and D.M. Magee, 2004. Coccidioidomycosis: host response and vaccine development. Clinical Microbiology Reviews, 17:804–839.

Elliot, P., Wakefield, J.C., Best, N.G., and Briggs, D.J., Spatial epidemiology: methods and applications, 2000, Oxford University Press, Oxford U.K.

Emmons, C.W., 1942. Isolation of Coccidioides from soil and rodents. Public Health Reports 57(4): 109-111.

Fisher, M.C., Koenig, G.L., White, T.J., and J.T. Taylor, 2002. Molecular and phenotypic description of Coccidioides posadasii sp. nov., previously recognized as the non-California population of Coccidioides immitis. Mycologia 94: 73–84.

Fisher, F.S., Bultman, M.W., Johnson, S.M., Pappagianis, D., and E. Zaborsky, 2007. Coccidioides niches and habitat parameters in the southwestern United States. Annals of the New York Academy of Sciences, 1111: 47–72.

Galgiani, J.N., 1999. Coccidioidomycosis: a regional disease of national importance, Rethinking approaches for control, Annals of Internal Medicine, 130(4): 293-300.

Hector, R.F., Rutherford, G.W., Tsang, C.A., Erhart, L.M., McCotter, O., Komatsu, K., Anderson, S.M., Tabnak, F., Vugia, D.J., Yang, Y., and J.N. Galgiani, 2011. Public health impact of coccidioidomycosis in California and Arizona. International Journal of Environmental Research and Public Health, 8: 1150–1173.

Hoffmeister, D.F., 1986. Mammals of Arizona, The University of Arizona Press, Tucson AZ, 600 p.

Kolivras, K.N., Johnson, P.S., Comrie, A.C., and S.R. Yool, 2001. Environmental variability and coccidioidomycosis (valley fever), Aerobiologia 17: 31-42.

Komatsu, K., Vaz, V., McRill, C., Colman, T., Comrie, A., Sigel, K., Clark, T., Phelan, M., Hajjeh, R. and B. Park, 2003. Increase in coccidioidomycosis – Arizona, 1998-2001, MMWR 52(6): 109-112.

Lacy, G.H. and F.E. Swatek, 1974. Soil ecology of Coccidioides immitis at Amerindian middens in California. Applied Microbiology. 27:379–388.

Litva, A. and J. Eyles, 1995. Coming out: exposing social theory in medical geography, Health & Place, 1(1): 5-14.

Maddy, K,T., 1965. Observations on Coccidioides immitis found growing naturally in soil. Arizona Medicine 22: 281–288.

Meade, M.S. and R.J. Earickson, 2000. Medical geography, Second Edition., The Guilford Press, New York, NY, p. 14.

Merriam, C.H. and L. Steineger, 1890. Results of a biological survey of the San Francisco mountain region and the desert of the Little Colorado, Arizona. North American Fauna Report 3. U.S. Department of Agriculture, Division of Ornithology and Mammalia, Washington, D.C., 136 pp.

NASA, 2011. The Thematic Mapper, National Aeronautics and Space Administration, Washington D.C. http://landsat.gsfc.nasa.gov/about/tm.html.

Nguyen, C., Barker, B.M., Hoover, S., Nix, D.E., Ampel, N.M., Frelinger, J.A., Orbach, M.J., and J.N. Galgiani, 2013. Recent advances in our understanding of the environmental, epidemiological, immunological, and clinical dimensions of coccidioidomycosis. Clinical Microbiology Review 26(3): 505-525.

Pappagianis D., 1994. Marked increase in cases of coccidioidomycosis in California: 1991, 1992, and 1993. Clinical Infectious Disease, 19:S14 –S18.

Park, B.J., Sigel, K., Vaz, V., Komatsu, K., McRill, C., Phelan, M., Colman, T., Comrie, A.C., Warnock, D.W., Galgiani, J.N., and R.A. Hajjeh, 2005. An epidemic of coccidioidomycosis in Arizona associated with climatic changes, 1998-2001. The Journal of Infectious Diseases, 191: 1981-1987.

Philo, C., 2009. Health and Health Care. In Gregory, D.; Johnston, R.; Pratt, G., Watts, M., and S. Whatmore, (Eds.), The dictionary of human geography (Fifth Edition), Blackwell Publishing, Malden MA, ISBN 9781405132879

Schneider, E., Hajjeh, R.A., Spiegel, R.A., Jibson, R.W., Harp, E.L., Marshall, G.A., et al., 1997. A coccidioidomycosis outbreak following the Northridge, Calif, earthquake, The Journal of the American Medical Association, 277(11): 904-908.

Stacy, P.K.R., Comrie, A.C., and S.R. Yool, 2012. Modeling Valley Fever incidence in Arizona using a satellite-derived soil moisture proxy. GIScience & Remote Sensing 49(2): 299-316.

Stewart, R.A. and K.F. Meyer, 1932. Isolation of Coccidioides immitis (Stiles) from soil. Proceedings of the Society for Experimental Biology and Medicine, 29: 937-938.

Sunenshine, R.H., Anderson, S., Erhart, L., Vossbrink, A., Kelly, P.C., et al., 2007. Public health surveillance for coccidioidomycosis in Arizona. Annals of the New York Academy of Sciences, 1111: 96–102.

Swatek, F.E., 1970. Ecology of Coccidioides immitis, Mycopathotogia et Mycologia applicata, 40(1-2): 3-12.

Tamerius, J.D., and A.C. Comrie, 2011. Coccidioidomycosis incidence in Arizona predicted by seasonal precipitation, PloS ONE 6(6): e21009, doi:10.1371/journal.pone.0021009.

USGS, 2013. The National Elevation Dataset (NED), U.S. Department of Interior, U.S. Geological Survey, http://www.usgs.gov/pubprod/data.html#data.

WHO, 2006. World Health Organization, Constitution of the World Health Organization – Basic Documents, Forty-fifth edition, Supplement, October 2006.

APPENDIX A

NOCTURNAL RODENT ABUNDANCE DISTRIBUTION IN SOUTHWESTERN ARIZONA ESTIMATED BY REGRESSION METHODS WITH REMOTE SENSING-DERIVED DATA

(TO BE SUBMITTED TO *ECOLOGICAL MODELLING*)

ABSTRACT

Nocturnal rodents play a key role in the Sonoran Desert ecosystem as consumers and prey. They interact substantially with the soil substrate and alter soil properties. They are also linked to the transmission of human diseases. Rodent distributions are poorly mapped for the study area, and factors derived from remotely-sensed data that explain rodent biogeography are not known. We build automated multi-linear regression models and decision tree models for ten rodent species using rodent trapping data from the Organ Pipe Cactus National Monument (ORPI) in southwest Arizona with a combination of surface temperature, a vegetation index and its texture, and a suite of topographic rasters as explanatory variables. The parametric and non-parametric models show remarkable congruency between themselves and with outside sources. Surface temperature is the most widely selected explanatory variable in both automated methods; we attribute this to its integrating capabilities of multiple landscape characteristics. Model output is used to develop spatially explicit distribution maps.

Keywords: nocturnal rodents, desert rodents, pocket mouse, kangaroo rat, wood rat, cactus mouse, abundance, remotes sensing, Landsat, GIS, classification and regression trees, decision Trees, stepwise regression, surface temperature, thermal infrared, vegetation index, topography, Sonoran Desert, Organ Pipe Cactus National Monument, Valley Fever, biogeography, ecological modeling, species distribution.

1. Introduction and Background

1.1 Introduction

Rodents are common in almost all habitats. They constitute a key prey base for carnivorous animals. They are responsible for considerable soil excavation, soil mixing, predation on plants and seeds, and dispersal of seeds (Reichman and Price, 1993). They thus alter the physical and chemical properties of soil. Rodents are implicated as carriers or hosts of several pathogenic agents causing human disease, including hantavirus, plague and coccidioidomycosis (Childs et al., 1994; Boone et al., 1998; Gage and Kosoy, 2005; Emmons, 1942; Kolivras and Comrie, 2004).

Nocturnal desert rodents are easily captured and identified, have small home ranges, have high fecundity, and respond quickly to changes in primary productivity and disturbance (Petryszyn, 1995). They provide a most cost-effective indicator for monitoring the primary consumer component of ecosystems (Holm, 2006)

Recent decades have seen an explosion of interest in species distribution modeling (SDM; Franklin and Miller, 2009). This has resulted from a confluence of the growing need for information on the geographical distribution of biodiversity, and of new and improved techniques and data suitable for addressing this information, such as remote sensing, global positioning system technology, geographical information systems (GIS) and statistical learning methods (Franklin and Miller, 2009).

We aim to characterize and predict the spatial dimensions of nocturnal desert rodent species and groupings throughout the Organ Pipe Cactus National Monument (ORPI), in southwestern Arizona, and beyond, using statistical modeling of data derived from remote sensing and GIS. In the first section of this article, we introduce the ORPI's

biogeography within the greater Sonoran Desert, and we identify common nocturnal desert rodent species that inhabit ORPI. We summarize current ecological modeling methods with a focus on stepwise regression and regression trees.

We explore in section two a rich source of rodent trapping data that spans over fourteen years and thirty trapping sites throughout the monument. We also describe spatial data sets, produced with remote sensing and GIS, that characterize the biophysical factors at play in determining rodent abundance, and we identify a concise set of these to build models.

In section three, we use parametric stepwise multiple regression (SWR) and non-parametric classification and regression tree analysis (CART), along with GIS techniques, to arrive at congruent, simple models to predict rodent abundance. We describe results in section four in the context of variable multicollinearity and spatial autocorrelation.

We conclude in section five with a comparison of model performance, and we address their limitations. We also suggest interpolation and extrapolation to produce maps of rodent abundance in ORPI and analogous regions throughout southern Arizona and in the Sonoran Desert.

1.2 Study Area

1.2.1 Sonoran desert

The Sonoran Desert stretches approximately 260,000 km^2 in southwestern Arizona, southeastern California, most of Baja California, the islands of the Gulf of California, and the western half of the state of Sonora, Mexico. The Sonoran Desert is

characterized for its semi-arid climate, mild winters and a bimodal rainfall pattern. Mountain ranges, a few volcanic hills, bajadas (coalesced alluvial fans), valley floors, and drainage systems (canyons, arroyos (steep-sided gulches), washes, and sheet-flow areas) are typical of its geomorphology (Hoffmeister, 1986). The visually dominant elements of the landscape are two characteristic life forms: legume trees and large columnar cacti (Arizona-Sonora Desert Museum, 2013).

The Lower Sonoran Life Zone (LSLZ), based on Merriam's elevation-precipitation life zones for the southwest, characterizes the lower elevation areas in the Sonoran Desert (generally > 1000m); these correspond to the warmer areas receiving lower precipitation (generally < 250mm), and populated with shrubs and succulents (Merriam and Steineger, 1890). The LSLZ is distinguished from the Upper Sonoran Life Zone, which characterizes upper elevations including mountains in the region.

1.2.2 Organ Pipe Cactus National Monument

The Organ Pipe Cactus National Monument (ORPI) was established in 1937 under the Antiquities Act in order to preserve a representative area of the Sonoran Desert (ORPI, 2006). The monument is situated in southwestern Arizona, near the geographical center of the Sonoran Desert (Figure 1). Its southern border lies along the Mexican border and adjacent to the Mexican state of Sonora.

Its namesake, the organ pipe cactus (Stenocereus thurberi) is a species of columnar cactus with multiple stems. The monument encompasses 133,882 hectares, and 95% is designated as wilderness. It is one of the most biologically diverse protected areas within the Sonoran desert biome in either the United States or Mexico (ORPI, 2006).

The city of Ajo, AZ (population 3,705 in 2000; Census, 2000) is situated approximately 20 km north of the park boundary, and the city of Senoyta, Sonora (Mexico; population 16,500 in 2000; ORPI, 2006) is situated approximately 5 km south of the park boundary (Figure 2). Arizona's largest cities, Tucson and Phoenix, are located approximately 240 km east and 200 km north, respectively. Arizona State Route 85, which connects these cities to beaches and resorts of Puerto Penasco, bisects the monument.

The monument is not immune to impacts from human activities. Highway maintenance activities and high traffic volume affect the natural resources along the right-of-way. Undocumented aliens migrating north from Mexico are deleteriously affecting the park resources (ORPI, 2006). Adjacent land uses include grazing, agricultural and residential development, particularly around the international boundary. Associated threats posed by these include compaction of the soil, trampling of vegetation, accumulation of wastes, groundwater overdraft, spread of exotic species, wildfires, pesticide drift, and feral or escaped domestic animals, woodcutting, and poaching.

1.3 Physiography of ORPI and the region

1.3.1 Climate

Summer months in the region are typically very hot with temperatures commonly above 100°F. Nighttime temperatures differ greatly from daytime (Hoffmeister, 1986) due to, in part, lack of moisture in the air and the absence of cloud cover. Cool summer evenings permit nocturnal animals in the desert to forage widely without great loss of

water. Daytime winter temperatures are mild, while nighttime temperatures are much cooler. Most of the area rarely experiences frost.

The region's precipitation regime is bimodal. The North American monsoon, which dominates the summer, is driven by seasonal high pressure system over the region with concurrent change in wind direction and advection of moisture from the Gulf of Mexico, the Gulf of California and the Pacific. Diurnal heating over desert surfaces creates hot, rising, moist air which cools and condenses in the upper atmosphere forming convective clouds and precipitation. The resultant spontaneous thunderstorms are accompanied heavy runoff.

Winter conditions are dominated by high pressure where anomalous conditions linked to Pacific/North American teleconnection pattern (PNA) and southwest troughing steer storms into the area. The resultant gentle and widespread rain may last for one or two days. The fore-summer and fall are typically dry. Arizona climate, particularly winter precipitation, is associated with the El Niño Southern Oscillation, and thus linked to global patterns.

Mean annual precipitation measured at a series of weather stations throughout ORPI is 240.5 mm (9.468 inches, 1943-2005; ORPI, 2006). The average annual temperature for this period ranged from 19.8C (67.7°F) to 22.6C (72.7°F), and the number of days with freezing temperatures ranged from 34 occurring in 1949 to 2 occurring in 1995.

1.3.2 Topography

Topographic features in ORPI are characteristic of the Basin and Range Physiographic Province (Warren et al., 1981). Steep, rugged mountain ranges formed by block-faulting, trend north-south. These are separated by broad, structural valleys.

The Ajo Mountains form the eastern boundary, including Mount Ajo (1465 m). Alluvial fans spread outward from mountain canyons, coalescing to form gently sloping depositional aprons (bajadas). Alluvial plains account for approximately two-thirds of the monument area (Warren et al., 1981).

The desert plains (Valley of the Ajo and Senoyta Valley) to the west of the Ajo Mountains average approximately 485m elevation. These plains, approximately 12 km in width, are gentle in slope and fan out broadly until interrupted by the Growler, Bates and Puerto Blanco Mountains, which rise up to 975 m in the center of the monument. To the west and north of these mountains stretches the Growler Valley, with an average elevation of 365m. The Cipriano and Quitobaquito Hills (450 to 600 m in elevation) are located in the southwestern portion of the monument. The lowest elevations, approximately 300 m above sea level, are found in southwest corner of the monument.

1.3.3 Geology and soils

The rock formations in ORPI are generally the result of volcanic activity and associated intrusions (Warren et al., 1981). Faulting in the late Tertiary age led to the upward displacement of these materials, which have since eroded to form ORPI's present day mountains. Ancient volcanic rocks including tertiary granites, basaltic andesite, basalt, schists and gneisses are exposed in many of its hills and mountains. The only

well-bedded sedimentary rocks, occurring along the northern portion of the monument, are shale, sandstone, and limestone and its metamorphosed conglomerates.

ORPI soils are developed from the weathering of volcanic rock. Prominent features of desert soils are their low organic content and poor development of soil horizons (Warren et al., 1981). Soil formation is slow, as chemical decomposition is impeded by the lack of rainfall. Organic matter formation is minimal due to slow production of humus, reduced oxidation due to extreme heat, and transport of detritus to drainageways by occasional intense rains.

The action of heavy rainfall through sparse vegetation and ephemeral streams produces a downgrade movement of alluvial material, with lighter material carried further. Erosion of this kind produces bajadas with characteristic, uniform soil texture gradient. Finer soil particles at edges and valleys occur as silts and clay and often produce loamy and sandy soils with little or no soil horizon development. Soils on piedmonts typically are shallow, with bedrock protruding, and often derived from a single parent material.

Infiltration rates generally vary with particle size; lower ORPI areas with finer textured soils have slower infiltration rates, greater runoff and lower amounts of available moisture. Small differences of available moisture may have profound effects on diversity and density of vegetation. Caliche (precipitated carbonates) and desert pavement (closely-packed pebble layer with deflocculated soil colloids), are common features of ORPI soils, and they also result in lower water infiltration rates and a consequent bareness of vegetation (Musick, 1975). They often impede rodent burrowing.

All of the soils in the monument are grouped in associations of very hot arid soils (aridosols). Sixteen soil series and one upland rocky association have been identified in the monument (Chamberlain, 1972). Deep, gravelly, calcareous soils on upper slopes account for 38.8% of the ORPI area (Warren et al., 1981); very shallow soils on low hills and mountains comprise 21.4%; deep soils on floodplains and alluvial fans account for 18.1%; rocky, stony, and rock outcrops comprise 17.4%; recently deposited alluvium along eroded stream channels comprise 3.8%, and deep soils on valley plains and terraces comprise 0.5% of area.

1.3.4 Vegetation

Desert scrub vegetation accounts for approximately 95% of the area (Warren et al., 1981). By far, the two most dominant communities are the paloverde-mixed cacti community (Arizona Upland phytographic division of the Sonoran Desert); and the creosote-bursage community (Lower Colorado subdivision). These two vegetation groups are described by Shreve and Wiggins (1964) as the only two major Sonoran Desert subdivisions to occur in Arizona; the other three only occur in Mexico. Corresponding Regional GAP (ReGAP, 2006) classes are the Sonoran Paloverde-Mixed Cacti Desert Scrub and Sonora-Mojave Creosotebush-White Bursage Desert Scrub (Figure 3).

Creosotebush associations are found primarily on nearly-level, fine textured, silty, permeable, and alkaline soils of the valley bottoms. They are distributed mainly in the western portions of the monument in the drainage of the lower Colorado River

(Steenbergh and Warren, 1977), and in the central plains. Additional plants are bursage, saltbush, desert thorn, mesquite, and galleta.

Paloverde-mixed cacti associations occupy the rolling upper bajadas and the lower mountain slopes where coarser soil texture and increased gravel and rock content provides greater available soil moisture for plants (Yank and Lowe, 1956; Bingham, 1963). This is a cactus desert with bursage, creosote bush and palo verde on the flanks of and extending over many desert ranges. Large cacti, saguaro, and organpipe may be present.

The primary environmental gradient occurring on a wide scale at ORPI progresses in southwest-northeast direction along the Colorado River basin, and affects a variety of factors including elevation, rainfall, temperature and soil characteristics. Smaller-scale environmental-vegetation gradients include the transition in plant-available soil moisture from the species-poor creosotebush associations on the lower bajada to the floristically and structurally diverse paloverde-mixed cacti associations on the upper bajada (Yang and Lowe, 1956), and the transition associated with elevation from desert scrub to woodland on mountains.

1.4 Nocturnal desert rodents

1.4.1 Heteromyid and murid rodents

Two families of rodents (Order Rodentia) dominate the nocturnal rodent communities of the Sonoran Desert, the Heteromyidae (heteromyids) and the Muridae (murids). The heteromyids are all nocturnal, burrowing animals with external fur-lined cheek patches for storing and transporting their primary food source, seeds. They are

well adapted to the arid environment and most do not need to drink water. The murids, or "true" rats and mice, are complex grouping of rodents, and various classification schemes differ on their treatment. Hoffmeister (1986) provides a thorough description of the classification of the rodents in Arizona, summarized here.

Two genera, the pocket mice (*Chaetodipus* spp. and *Perognathus* spp.) and the kangaroo rats (*Dipodomys* spp.) represent the heteromyid family in Arizona. Six heteromyids, including two kangaroo rats and four pocket mice, are commonly found in ORPI during annual trapping events (ORPI, 2006).

The murid family includes the subfamily Cricetinae, or New World mice and rats. In some classification schemes, the "Cricetidae" are a distinct family. Several members of the cricetids, including the genera *Peromyscus* (white footed mice), *Neotoma* (wood or pack rats), *Onychomys* (grasshopper mice), and *Sigmodon* (cotton rats), are commonly found in the trapping study. Genera *Baiomys* (pygmy mice) and *Reithrodontomys* (harvest mice) are not commonly found.

1.4.2 Rodent adaptation to the desert

Small and large mammals live and do well in the arid, hot deserts of Arizona (Hoffmeister, 1986). Often there is little or no free water, little or no shade, extremely high temperatures during the daytime, limited amount of food, and low humidity. Rodents overcome these problems in physiological, morphological, and behavioral adaptations.

Kangaroo rats and pocket mice can exist with no free water and relatively dry food throughout their life. They manufacture water from the seeds they eat by oxidation,

conserve water they produce, and store seeds for use throughout the year. Efficient kidneys form urea four times more concentrated as humans, and waste products are almost dry. They are nocturnal, and thus they are not above the ground in the heat of the day.

Burrows are often one meter below the surface where daily temperature fluctuations are one fifth of that at the surface. The burrows leading to nests are often plugged with soil much of the time to prevent influx of heat and escape of moisture.

Desert vegetation produces large numbers of seeds, which may number in the thousands per square meter at certain seasons (Hoffmeister, 1986; Dye, 1969). Deserts seeds become widely distributed by wind and water, and often partially hidden by soil. Due to the unpredictability and the infrequent but intense production in short times, desert rodents store seeds for relatively long periods of times. Also, many desert rodents have cheek pouches to aid in gathering large quantities of seed. These can contain up to 900 seeds in the kangaroo rat.

1.5 Causal factors for species abundance

Austin (2002) distinguished direct (proximal or causal) factor gradients, or those with direct physiological effect (water, light, nutrients, prey, and nest site availability, for example) on species distribution. Indirect factor gradients, which have no direct effect, are always distal factors. Latitude, longitude, elevation, slope (steepness), and aspect (exposure) are examples of these. Distal factors are related to resources or regulators, and they can usually be correlated to species distributions (Franklin and Miller, 2009). They are often easier to measure or observe than direct factors. When only indirect factor

gradients are used, caution is advised to not extrapolate results beyond the range of conditions used to develop the model. Furthermore, there is no theoretical expectation for the shape of a species response curve on an indirect gradient (Austin, 2007).

Response curves, or resource selection functions, for animals depict a function describing the relationship of species abundance in relation to values of environmental (factor) gradients (Whittaker, 1967). Species response functions may be Gaussian, however unimodal, skewed responses to resource gradients are common (Franklin and Miller, 2009). Resource gradients with physiological stress can result in skewing at the harsh end, and competition can limit the benign end. Bimodal or multimodal model response curves have been hypothesized to result from competition (Whitaker, 1960).

The true response of a species to one resource factor is detectable only when all other factors occur at non-limiting levels, according to Liebig's law of minimum (Huston, 2002). In reality, it is not usually possible to observe this without experimental data, and factors tend to co vary. Quantile regression, modeling the upper and lower bounds has been suggested as an alternate approach to this problem (Huston, 2002).

Abundance is usually derived from numerical abundance, or number of individuals, because this is the currency in which the taxa (e.g. birds and trees) that predominate in such analyses are typically recorded (Henderson and Magurran, 2010). Biomass, as an alternative, is preferred by some ecologists (Saint-Germain et al. 2007). Body mass of individual can be correlated to metabolism scales, and it is assumed to provide a more direct measure of resource use (Kleiber 1962).

1.6 Ecological modeling methods

1.6.1 Statistical modeling

A strong distinction is made between general ecological models (conceptual or heuristic) and a distinct subset of these, statistical models (Guisan et al., 2002). In most studies, some sort of conceptual or theoretical model (Austin, 2002) of the ecological system is already proposed before a statistical model is considered (Guisan & Zimmermann 2000). The statistical model provides a mathematical basis for interpretation by examining parameters such as fit, or how well the measured predictors explain the response, strength of association, and the contributions and roles of the different variables (Zar 1999; Legendre & Legendre 1998).

Statistical models in ecology are further distinguished by goal. Explanatory models seek to provide insights into the ecological processes that produce patterns (Austin et al. 1990), such as to ascertain the strength of the statistical relationship between a response. In contrast, predictive models typically seek to provide a statistical relationship between the response and a series of predictor variables for use in predicting the probability of species occurrence or estimating numbers of an organism at new, unsampled locations. These models often use variable reduction techniques in order to predict the ecological attributes of interest from a restricted number of predictors. The concept of parsimony, that the simplest explanation is best, is inherent in such modeling efforts.

Numerous modeling approaches are reported for wildlife distribution modeling, including Generalized Linear Models (GLMs), Generalized Additive Models (GAMs), Multivariate Adaptive Regression Splines (MARS), Genetic Algorithm for Rule Set

Production (GARP), decision trees, machine learning models and model averaging methods such as bagged or boosted classification trees. Franklin and Miller (2009), provide a summary of many of these methods.

Inter-model comparisons have yielded conflicting results about the relative performances of different models (Franklin and Miller, 2009; Dormann et al. 2008). For example, single CART analysis is reported to have lower accuracy and underestimates species distribution as it is prone to omission error; however GARP overestimates and is prone to commission error (Lawler et al., 2006). One approach to study design is to use more than one model method; this is likely to yield insights into the underlying species patterns as well as model structure and assumptions (Franklin and Miller, 2009).

1.6.2 Linear regression and stepwise techniques

Linear regression is one of the oldest statistical techniques used in ecological research. It is a global method, i.e. it uses all the data, to estimate a linear relationship (Guisan et al., 2002). The basic linear regression model has the form:

$$Y = \alpha + X^T\beta + \varepsilon,$$

where Y denotes the response variable, α is a constant called the intercept, $X = (X_1, ..., X_p)$ is a vector of p predictor variables, $\beta = \{\beta_1, ..., \beta_p\}$ is the vector of p regression coefficients, one for each predictor, and ε is the error. The error represents measurement error, as well as any variation unexplained by the linear model, which one tries to minimize when fitting the model. Implicit in the application of regression tools for species modeling is a pseudo-equilibrium between the organisms and their environments (Austin, 2002).

Linear regression is limited by three main assumptions about the data: (1) the errors ε_i are assumed to be identically and independently distributed; this includes the assumption that the variance of Y is constant across observations; (2) the errors ε_i are assumed to follow a normal (Gaussian) distribution; and (3) the regression function is linear in the predictors. The end result is typically stable but possibly inaccurate predictions with low variance and high bias. Additional challenges in regression modeling are variable selection, multicollinearity in the response variables, and interactions between the predictors.

Common ways of addressing these limitations are transformations of the response variable to meet the criteria of normality and constant variance, such as the Box-Cox approach (Mateu 1997), or augmenting the predictors with polynomial terms, interactions and other non-linear transformations in order to build a non-linear model in terms of the Xj but linear in the parameters (Guisan et al., 2002).

Evaluation criteria, such as deviance reduction measured with the chi-statistic or the Akaike Information Criterion (AIC, Akaike 1973; Sakamoto et al. 1988) can assist in variable selection. Collinearity can be detected with approaches such as a condition number and a variance inflation factor (VIF; Brauner & Schacham 1998). The use of classification and regression tree (CART) as a complement to regression models is a promising approach to identify these limitations (Hastie et al., 2009). Despite these challenges, regression analyses have been broadly applied in SDM (Guisan and Zimmermann 2000; Scott et al. 2002).

1.6.3 Automated regression techniques

Automated model selection with linear regression can employ backwards elimination, forward selection, or stepwise (uses both forward and backward). Stepwise regression (SWR) is a systematic method for adding and removing terms from a multilinear model based on their statistical significance (MATLAB, 2013).

SWR begins with an initial model and then compares the explanatory power of incrementally larger and smaller models. The method proceeds as follows: 1) fit the initial model; 2) if any terms not in the model have p-values less than an entrance tolerance (that is, if it is unlikely that they would have zero coefficient if added to the model), add the one with the smallest p-value and repeat this step; otherwise, go to step; 3) if any terms in the model have p-values greater than an exit tolerance (that is, if it is unlikely that the hypothesis of a zero coefficient can be rejected), remove the one with the largest p-value and go to step 2); otherwise, end. The method terminates when no single step improves the model.

Stepwise procedures are considered to be high-variance operations because small perturbations of the response data can sometimes lead to vastly different subsets of the variables. They should be used with care (Guisan and Zimmermann 2000). They can be improved if selection criteria based on permutations of the data, such as 10-fold cross-validation, are used (Hastie et.al 2009).

1.6.4 Classification and Regression Trees (CART)

Decision trees (DT) are divisive, monothetic, supervised classifiers (Franklin et al. 2006). The DT is as set of nested, binary decision rules that is used to classify data

into homogenous subgroups (nodes) based on threshold values of the predictors (Hastie et al., 2001). The splitting rules define the branching at the internal nodes. If the dependent variable is categorical or ordinal, the result is a classification tree; if the dependent variable is continuous, the result is a regression tree. The predicted value for the regression tree case is the average value of the training data in that node.

The DT process takes place in three stages: tree growing, tree stopping and tree pruning. The basic approach in DT production has been to "grow" a large DT with fairly liberal stopping rules and then to "prune" the tree, i.e. remove splits or collapse the internal nodes that add least to overall subgroup homogeneity. The final size (number of terminal nodes) is usually determined by comparisons of overall tree deviance (sum of residual sum of squares in the nodes) or by some form of k-fold cross validation procedure.

DTs often require larger samples than other methods because successive splits are based on fewer and fewer observations (Vayssieres et al., 2000). Trees can be unstable. For example, varying the inputs, such as sampling from a different set of observations, or varying the set of explanatory variables, can result in very different models (Hastie et al., 2009; Scull et al., 2005). Also, it is difficult to interpret variable importance. Finally, situations with a continuous, linear species response may not be well characterized with a method based on threshold rules.

DTs often provide an alternative modeling method for comparison with traditional methods (Franklin and Miller, 2009). They are particularly useful at categorical predictors, hierarchical interactions, and data with threshold characteristics. They provide an output that is visual. While regression examples of DTs are relatively

uncommon, they are expected be useful in ecology (Franklin and Miller, 2009), especially with continuous data sets derived from remote sensing and GIS. One approach is to approximate a linear response with step functions (De'ath and Fabricious, 2000). New methods employing ensemble modeling (model averaging), such as bagging, boosting, and random forests are beginning to find success in SDM modeling (Franklin and Miller, 2009).

1.7 Remote sensing and Geographical Information Systems in species modeling

Remote sensing's importance in geological and environmental sciences is unsurpassed (Vincent, 1997). Remote sensing provides a synoptic view of the landscape, and it is a source of multispectral data. Data is obtained at a variety of spatial resolutions and extents, and repeat visits provide a source of temporal assessments. Integrated remotes sensing and Geographical Information System (GIS) provide a platform for spatial analysis. These provide alternatives to traditional ground surveys for the production of species distribution maps.

Spatial prediction in SDM is contingent upon the availability of environmental predictors in the form of maps, or more precisely, digital spatial data (Goodchild, 1996). For a growing number of places in the world, there are many available geospatial and remote sensing data sets, These are often obtained free from public agencies. However, the suitability of spatial or (GIS) environmental data sources for analytical use in ecological modeling has not been given as much attention as other aspects of SDM (species data, modeling methods, etc.; Hunsaker et al., 1993). The challenge is to identify environmental maps that represent resource gradients or other factors determining species

distribution at the appropriate scale. This effort involved is non-trivial, and comprehensive guidelines are not available (Franklin and Miller, 2009).

Virtually all SDM studies at global and regional scales use a climate parameter; modeling at smaller extents generally use topography as surrogates for temperature and moisture gradients (Franklin and Miller, 2009; Wilson and Gallant, 2000). Topographical variables used in SDM can be related to direct and resource gradients; for example, the effect of elevation on temperature and precipitation, of slope on radiation regime and moisture availability, and of landform, hillslope positions and catchment position on soil moisture, erosion and deposition, which in turn affect soil development and properties (Franklin, 1995).

Wildlife habitat suitability models also traditionally rely on vegetation categories as one of their main drivers (Kochler and Zonneveld, 1998). However, vegetation categories mapped for multiple land management purposes may not be a suitable match to describe habitat suitable for a target species (Franklin and Miller, 2009). Some aspect of structure such as cover, size distribution of plants, vertical canopy structure, may be also related to habitat suitability. Despite these challenges, many researchers have used vegetation maps, including the Normalized Difference Vegetation Index (NDVI), to predict species abundance and richness (Oindo and Skidmore, 2002).

2. Data

2.1 Rodent trapping data

We obtained rodent trapping data from Organ Pipe Cactus National Monument by information request. ORPI staff and associated researchers monitor rodents annually over two consecutive nights in the summer (June-August) at 35 trapping plots (grids)

since 1991. They use a consistent field study protocol based on Petryszyn (1995). Trapping methods and sampling plans for this data set are described in more detail in ORPI (2006).

We truncated the data set to cover the years 1993 through 2006 (fourteen years) and 30 trapping sites in order to provide fullest coverage of years for all sites, and we omitted five plots (BURN1, BURN2, CAMP1, LOST1 and LOST2) in order to remove data with the most gaps. Nineteen of the 30 remaining trapping plots had complete coverage for truncated period. The remaining 11 plots were missing between one to four years of data; we included data from these sites in order to avoid decreasing the number and diversity of observations.

For each plot, we calculated a relative annual abundance index using the count data and adjusting for effort on (ORPI, 2006; Beauvais and Buskirk, 1999) as follows: N_r = [Total unique individuals caught/((traps*2 nights) – (0.5 * sprung traps))] * 100%.

We tabulated a relative annual abundance value for each species at each plot, and we determined means of the annual values to represent the fourteen year study period. We standardized or z-scored the mean fourteen year values for each species across all 30 sites, using z-score value = (mean – observation)/standard deviation. Figure 4 provides the z-scored abundance data for the ten rodent species and their ranges.

We also developed biomass data using log mean capture weights reported in the trapping data set (ORPI, 2006). Finally, we determined diversity (Shannon's Diversity Index, or

$$H' = -\Sigma pi*\ln(pi)$$

where p_i is the fraction of the entire population made up of species *i*, and richness (number of species) indicators using the total capture data. Mean annual values for these were prepared as described above.

2.2 Reflectance image processing

The Landsat Thematic Mapper 5 (TM) provides reflective data since 1984 in six spectral regions, including the visible, near- and mid-infrared, at 30 meter spatial resolution (NASA, 2012). The satellite also provides data in the thermal infrared with one band at 120m spatial resolution. Images are available for free access from the United States Geological Survey (USGS) Earth Explorer (http://earthexplorer.usgs.gov).

ORPI is located in Path 37 and Row 38 of the Landsat Worldwide Reference System. Landsat Thematic Mapper 5 (TM) archive images are processed to the Level 1 Product Generation System with precision and terrain correction (LPGS, L1T), and with cubic convolution resampling (USGS, 2011).

To minimize annual phenological effects, we selected Landsat 5 TM images captured in June for each year of the study period (1993-2006). June images also match the rodent trapping periods. We incorporated updated radiometric calibration coefficients specific to the new USGS Landsat open-access archive (Chander et al, 2009) to achieve conversion of calibrated digital numbers to absolute units of at-sensor spectral radiance. We converted each image to atmospheric-corrected surface reflectance using the cosine approximation model (COST; Chavez, 1996). The COST model implements an improved dark-object atmospheric correction for Landsat TM multispectral data (bands 1-5 and 7). Comparisons of random points within the area of interest showed good

spatial registration to within one half a pixel's dimension (15 m) between the images, and therefore additional fine-tuning beyond USGS rectification was not performed.

We composited the fourteen annual June images for each reflective band into a median value image for each pixel. We subsequently applied a 5x5 median spatial filter (low pass) to each 14-year median band image. The 5x5 filter window dimension was chosen to best capture the dimensions of a typical ORPI rodent plot (approximately 0.8 ha as compared to 2.25 ha filter window). A comparison of plot sizes and 5x5 filter windows on a vegetative index image for the LOWE1 and LOWE2 plots is shown in Figure 5. Coordinates for each rodent plot's center was used to extract the image pixel value for the spectral indices used in the study.

We prepared a Soil Adjusted Vegetation Index (SAVI; Huete, 1988; Huete et al. 1992) from the spatially-filtered annual median red and near-infrared images. The SAVI, derived from the Normalized Difference Vegetation Index (NDVI), incorporates a canopy adjustment factor to minimize soil noise inherent in the NDVI by accounting for differential red and near-infrared extinction through the canopy. The SAVI image is generated using the following formula:

$$SAVI = (1+L)(\rho nir - \rho red)/ (\rho nir + \rho red + L),$$

where ρ is the respective spectral band reflectance, and L is the canopy background adjustment factor (0.5). SAVI values were rescaled to 0-1 to remove negative numbers.

We also generated a texture image from the SAVI using Erdas *Imagine* Image Interpreter based on a variance (2nd order) operator on a 5x5 moving window as follows:

$$Variance = \Sigma (x_{ij} - M)^2/(n-1),$$

where x_{ij} = DN of pixel (i,j), n = number of pixels in window, and M = mean of the moving window. Texture can provide information about local variability of pixel intensity values pixels in an image. For example, in areas with homogenous, smooth texture, the range of values in the neighborhood around a pixel will be smaller, resulting in a low texture value; in areas of high pixel spectral diversity, or rough texture, the range will be larger. Figure 6 provides statistics of the predictor variables including SAVI and its texture.

2.3 Topographical image processing

We prepared topographical images using Digital Elevation Model (DEM) data from the USGS National Elevation Dataset (NED, 2013; one arc second data, roughly 30 meters). We extracted and performed a mosaic image with eight DEM subpanels to generate elevation data covering the whole region.

Using the ESRI *ArcMap* Spatial Analyst Surface toolset, we generated slope (gradient, or rate of maximum change in z-value), aspect (slope direction of the maximum rate of change in the z-value from each cell in a raster surface), and curvature images (the second derivative value of the input surface on a cell-by-cell basis, or degree of concaveness or convexness). Mathematically, curvature is defined as the reciprocal of the radius of a circle that is tangent to a point on a curve. Tightly folded terrain has large curvature values, while flat terrain has zero curvature. We generated a "distance to wash" raster using the DEM with ESRI's ArcMAP stream network tools including Flow Accumulation, and we prepared easting and northing images from the aspect image using the functions below (Zar, 1999):

$E = \sin((A*\Pi)/180)$,

$N = \cos((A*\Pi)/180)$,

where E is easting, N is northing, and A is aspect.

We applied a 5x5 median filter to each topographical image as described above for the Landsat reflective images in order to achieve effective sampling of rodent plot values.

2.4 Surface temperature image processing

We prepared a surface temperature image for the study area using a radiative transfer model and surface emissivity estimated from a vegetation index (Sobrino et al., 2004; Barsi et al., 2005; Jiménez-Muñoz et al., 2009). We obtained annual Landsat TM thermal infrared images (band 6, 10.4-12.5 µm) captured in June dates for the period of 2000-2006. We did not use available thermal images for the years 1993-1999 because atmospheric correction data were not available for earlier years.

We rescaled the calibrated digital numbers of the L1 thermal band products to at-sensor spectral radiance using gain and offset coefficients provided in Chander et al. (2009). To estimate surface-leaving radiance for each pixel, we adjusted the at-sensor spectral radiance for atmospheric attenuation and enhancement using an inversion of the radiative transfer equation. Atmospheric constants necessary for this were obtained for each scene from the Atmospheric Correction Parameter Calculator (Barsi et al. 2005; http://atmcorr.gsfc.nasa.gov), a web based tool which uses a MODTRAN radiative transfer model plus date, time and location to provide upwelling (atmospheric path)

radiance, downwelling (sky) radiance, and atmospheric transmission factors specific to Landsat systems thermal bands.

We estimated surface temperature for each pixel using its respective surface-leaving radiance (above) and its emissivity. We estimated emissivity with the simplified normalized difference vegetation method (Jiménez-Muñoz et al., 2009; Sobrino et al., 2008), which required preparation of a fractional vegetation component (FVC) image using a normalized difference vegetation image of the area. Inputs to the FVC calculation are the $NDVI_{soil}$ and $NDVI_{vegetation}$, which we extracted from histograms of scene NDVI, and estimated to be 0.05 and 0.75, respectively. We used a linear FVC equation, which we found, in contrast a nonlinear function used in Jiménez-Muñoz (2009), was a better representation of the ORPI desert landscape. A linear function is supported in other literature (Montandon and Small, 2008; Xioa and Moody, 2005; Zeng et al., 2000).

Additional inputs to the emissivity calculation include emissivities of vegetation and of soil components of any one pixel for the ORPI scene. We estimated these at 0.982 and 0.958, respectively for the ORPI area using the North American ASTER Land Surface Emissivity Database (NAALSED, 2013; Hulley and Hook, 2009) with a similar spectral band pass. These estimates compared well with literature estimates of similar terrain (Sanchez et al. 2011; Schott; Humes et al., 1994; Sobrino et al. 2004; Jiménez-Muñoz et al., 2009).

We prepared a median surface temperature image, similar to the reflective products, using the seven annual surface temperature images in a similar manner as

described in the reflective image processing. A surface temperature value for each plot was sampled using the plot's center.

A summary of surface temperature processing steps is presented in Appendix 1. The surface temperature image preparation entailed a significant amount of calculation and parameter estimation (soil and vegetation NDVI and emissivities, upwelling and downwelling radiances, etc.). The end result was an image with very high correlation to the original unprocessed Landsat TM band 6 digital number image (correlation coefficient of 0.979, R^2= 0.959; *p*-val=5.71E(-21)). We suggest that the unprocessed thermal image would be sufficient for use as a substitute for surface temperature in this Sonoran desert study.

2.5 Summary of data set scale and extent

The greater ORPI area constitutes the extent of the study area. The sample size is 30 rodent plots, all of which lie within ORPI boundaries. Each rodent plot is approximately 8100 m^2. A small sample size and plot area compared to the ORPI extent results in a very low sample density (approximately 1.8E(-4)). The prevalence or frequency of plots with a particular species, ranges from 0.057 (SIAR) to 1.00 (CHBA and CHPE), with an average prevalence of 0.67 for all ten species.

Abundance modeling literature supports 50-100 observations of species presence to produce acceptable results (Stockwell and Peterson, 2002; Loiselle et al., 2008), however studies with fewer have been shown to be successful (Elith et al. 2006, Wisz et al., 2008). The absolute number of observations, according to some authors (Franklin and Miller, 2009; Kadmon et al., 2003), may be less important than having observations

well distributed throughout the environmental space it occupies. To mitigate this, the rodent plots represent a suitable range of diversity found in ORPI (ORPI, 2006). This is further supported by histogram evaluations of exogenous variables. In addition, repetitive annual sampling for over fourteen years at each of the 30 plots is likely to significantly improve some aspects of statistical accuracy hindered by a limited number of observations.

The measurement scale in both the endogenous (abundance) and exogenous variables (surface temperature, vegetation index, topography) is continuous ratio type. Figures 7 and 8 depict histograms of DIME (kangaroo rat) and surface temperature, one of our exogenous variables (N=30 plots for both). Figure 9 shows the corresponding scatter plot relationship between DIME and surface temperature values. The histograms and scatter plots demonstrate sufficient characteristics needed to support continuous-scale models, such as spread and shape. Output of continuous, ratio type can be downgraded to ranked or classed data as needed.

3. Analysis

3.1 Model building

The composition of Lower Sonoran Life Zone relatively free from human disturbance is primarily soil, rock, and vegetation. We initially explored MODIS images and associated spectral indices in mid- and thermal infrared region to characterize the ORPI study area based on literature suggesting a soil response in the infrared spectrum (OMI reference, TIR-soil references). Despite initial success using MODIS imagery, we

determined that its pixel resolution of one kilometer was insufficient for the scale of measurement of the rodent monitoring data.

We therefore explored Landsat thermal infrared imagery, a finer resolution data set. Since ORPI is composed primarily of soil and vegetation, we added the vegetation index SAVI as a potential explanatory variable.

Despite success at characterizing some species with SAVI and surface temperature, we looked for additional variables to improve models and characterize other species abundance. Topographical data is also readily available through Digital Elevation Models, and it is directly associated with primary ecological resources, such as temperature, precipitation, and sunlight (Franklin and Miller, 2009). We added a suite of topographical variables to our models.

Adjusted coefficients of determination of various combinations of explanatory variables of our data set are shown in Figure 10. Models that failed to generate statistically significant results or failed other diagnostics such as multicollinearity or heteroscedacity are represented with a dash. The four columns on the right side show the progression of improvement in overall models, based on their correlation coefficients, as variable sets are added one at a time beginning with topographical variables. Correlation coefficients improve substantially upon addition of surface temperature to the topographical set of variables, and improve further upon addition of a vegetation index.

We also explored additional spectral indices generated from Landsat, including soil-focused ones such as a clay soils index and the iron oxide index. While these improved correlations coefficients, we did not include these in the final models due to a small increase in multicollinearity and challenges in explaining their results. We were

satisfied that the surface temperature, SAVI, texture, and topographical variables were sufficient to build models for most of the species.

3.2 Stepwise Regression

We applied MathWork's MATLAB stepwise regression functions *stepwise* and *stepwisefit* to our data set, using default entrance and exit tolerances for the *F*-statistic *p*-values (0.05 and 0.10, respectively). SWR generated models for nine out of the ten species. SWR failed to generate a model for SIAR, which had the most rodent plots with zero abundance. Figure 11 shows the SWR model for CHBA, the Bailey's Pocket Mouse as an example. In addition to models for individual species, we also estimated models for heteromyid and murid groupings and all species for abundance, biomass, and for Shannon's Biodiversity Index (SDI, diversity) and number of species (richness).

We applied regression diagnostics tests for multicollinearity (Variance Inflation Factor < 7.5), residual normality (Jarque-Bera *p*-value > 0.05), stationarity and residual homoscedasticity (Koenker's studentized Bresch-Pagan (BP) *p*-value > 0.05), and residual spatial autocorrelation (Global Moran's I *p*-value > 0.05) to the SWR models, and each species and groupings passed the diagnostic tests.

Stepwise models are locally optimal, and there is no guarantee that the produced model is globally optimal. To assess this, we performed SWR with predictor variables reversed in order for each species, and no changes in variable selection occurred. The data were also tested using exploratory regression for Akaike Information Criteria (AICc) values to check for alternative or superior models. Exploratory regression tools check all possible combination of predictor variables. In most cases, models predicted by stepwise

regression were those with the lowest AICc values. The SWR models for CHBA, CHPE and ONTO exhibited very similar AICc values to the best performing exploratory regression model. The SWR models were more parsimonious, and we present SWR results for all models in order to maintain consistent criteria for model selection.

Five species, CHIN, DISP, ONTO, PEER, and SIAR, exhibited a significant number of rodent plots with zero abundance. We removed the zero abundance values from the data sets before performing SWR for these species, and achieved improved statistical results. We therefore used their reduced- plot results for final model selection.

We tested non-linear exponential and power models for the ten species. The models for CHBA, CHPE, PEAM, DIME, and ONTO showed improvement in their coefficient of determination values. Other models, CHIN, NEAL, and PEER, did not. The average overall improvement in coefficient of determination for the ten species was 0.05. We report linear model results here for consistency and ease in interpretability. We also tested interaction terms for the explanatory variables with the SWR models, and we found no significant improvement of results.

3.3 Classification and Regression Tree (CART)

We performed regression tree analysis on the data set using MathWork's MATLAB classification and regression tree functions *classregtree(X,y)*. We adjusted MATLAB default parameters to account for the characteristics of the small-N data set (30 rodent plots). We decreased *minparent*, which is a number *k* such that impure nodes must have *k* or more observations to be split, from the default of 10 to 2. We left the value for *minleaf*, or the minimal number of observations per tree leaf, at its default value

(1). We increased *qetoler*, which defines the tolerance on quadratic error per node and determines when splitting is terminated, from 1E-6 to 1E-1. Splitting of a node is terminated when the quadratic error per node drops below *qetoler*qed*, where *qed* is the quadratic error for the entire data computed before the decision tree is grown. The adjusted parameters resulted in superior statistics, including lower deviances in terminal nodes and in the overall tree.

The original CART models were also analyzed for suitability for pruning based on cross-validation testing results. We used 10-fold cross validation testing on expanded tree sizes to assess the size and reliability of the default models. The function partitions the original sample into 10 subsamples, chosen randomly but with roughly equal size. For each subsample, the test fits a tree to the remaining data and uses it to predict the subsample. It pools the information from all subsamples to compute the cost for the whole sample. The cost of the tree is the sum over all terminal nodes of the estimated probability of a node times the cost of a node.

Plots of cross validation costs against node size were performed, and the tree with lowest cost was compared with that predicted by default model input parameters. For a majority of the ten species and groupings, the lowest-cost tree identified by cross-validation testing was the same as that predicted by our original adjusted default parameters. For CHIN, PEAM and DISP, pruning was performed to the smaller tree predicted by the cross-validation tests. For CHBA and DIME, pruning was not performed, as we observed undesired increases in standard deviations of resultant terminal nodes, which suggests that important relationships are potentially lost by pruning.

4. Results

4.1 SWR and CART results

Detailed statistics for the CHBA (Bailey's pocket mouse) SWR model are shown in Figure 11 as an example. Elevation is predicted to have the most effect on CHBA abundance with a coefficient of +0.567 (z-scored, 564 m actual value). Aspect-northing and distance to nearest wash are also positive predictors. The intercept constant is zero, and its p-value is one, due to z-scored input.

Figure 12 presents a summary of SWR models for the ten rodent species, groupings and diversity. Adjusted coefficient of determination and overall model *F*-statistic (*p*-value) are shown. Coefficient for the predictor variables are shown in order of magnitude, and negative coefficients are highlighted.

Figure 13 shows the predicted CART tree for CHBA as an illustration, and Figure 14 is a summary of CART results for CHBA. Abundance values are listed at the each of the five terminal nodes of the CART tree, and the terminal node statistics are shown in the boxes. A ranking of variable importance which governed CHBA's tree formation is shown in Figure 15. For CHBA, the root node or first split is based on aspect-northing, indicating northing is the primary exogenous variable operating at broad scales. The rodent plots, or observations, with aspect northing of less than 1.14637 (z-scored, or +0.3311 actual aspect value) are passed to the left, and those equal or greater are passed to the right. Both sides of the second-level splitting employ surface temperature. Texture of the SAVI image (diversity) is used as a final third level split. Terminal nodes are nodes 5 through 9. Highest CHBA abundance is predicted to be 2.43 at node 7,

composed of areas with aspect northing greater than 1.144 (+0.3311) and surface temperature greater than -0.717 (324.6 K).

Overall tree deviance plotted by tree size, or number of terminal nodes, for CHBA is shown in Figure16. Tree deviance is calculated as the sum residual squares (sum of terminal node sum of squares, observed versus mean) divided by the degrees of freedom (number of total observations minus number of terminal nodes). Note that the deviance drops drastically from a value of 1.0 in the root node (no splits) to approaching zero at five terminal nodes. By deviance criteria alone, the CHBA tree pruned to five terminal nodes sufficiently represents the data set. Figure 17 presents a plot of a cross-validation check for CHBA model reliability. Cross-validation cost testing for CHBA suggests that a tree pruned to four terminal nodes is best suited to minimize overall deviance, and larger tree sizes do not improve statistics adjusted for tree size.

Figure 18 lists summary CART statistics and main predictors for the rodent species, groupings and diversity, and Appendix X displays the respective CART trees.

4.2 Validation of results

An independent data set, or other trapping data for example, is not available for this area. One way to test the reliability of the SWR and CART models, without independent data, is by intra-data comparisons. Common approaches to this in SDM is either 1) partitioning the data set in a training set and a test group, such as k-fold cross-validating; or 2) bootstrap sampling (sampling with replacement) of the entire data set (Franklin and Miller, 2009; Guisan and Zimmermann, 2000; Wintle et al., 2005).

We tested the SWR models with a bootstrapping function based on 1000-fold sampling with replacement to test regression coefficient values. The variance on the output for each of the ten species was very low, and ranged from (0.0100 to 0.001). We therefore concluded the SWR models passed this check for validity.

Cross-validation testing used for tree pruning, described in Section 3.2, provide a means to perform inter-data validity checking for the CART models. For most of the ten species and groupings, successful cross-validation tests were accomplished, and the models passed this check for validity. For DISP and SIAR, the lowest cost cross-validation model was the unsplit root node. This is attributed to most of the observations, or rodent plots, for these two species were zero value, and therefore not enough diversity is expressed in the data set for model building.

4.3 Multicollinearity in the predictors

Correlation existed between the individual predictors. Correlation coefficients between surface temperature and the other variables are shown in Figure 19. The highest positive correlation coefficient for surface temperature is with curvature (+0.54), and the highest negative correlation is with slope (-0.73). Other high positive correlations include elevation-UTM_easting (+0.73, reflecting a general upward elevation gradient from the Colorado River basin), slope-elevation (+0.69), elevation-aspect_northing (+0.50), and SAVI-Text_SAVI (+0.50). High negative correlations include slope-curvature (-0.70), surface temperature-elevation (-0.48), and distance_wash-UTM_easting (-0.48). Despite elevated correlation coefficients between some predictors, regression diagnostics such as Variance Inflation Factor did not flag violations for models

produced in SWR. CART, a non-parametric method, is sensitive to multicollinearity among predictor variables.

4.4 Spatial autocorrelation

Global Moran's I coefficients for each of the ten species, determined by sampling their thirty-plot values using Euclidian distance and a default neighborhood distance of 8585 meters with no weighting, were statistically significant (except for CHIN and SIAR). The values ranged from 0.412 to 0.726. Moran's I values were determined also for the predictor variables, as sampled from the thirty rodent plots, and the values for the non-UTM variables ranged from 0.21 to 0.65. The highest was surface temperature (0.649, p-value 0.00292).

Significant spatial autocorrelation therefore is indicated in both exogenous and endogenous variables. Yet, testing with use of spatial lag and spatial error models did not significantly improve overall model performance based on correlation coefficients and statistical significance. The non-spatial regression models for all species passed diagnostic test for spatial autocorrelation (including statistically insignificant Global Moran's I for the model residuals, or p-value > 0.05).

We believe the explanation for this relates to the underlying structure of the landscape. Figure 20 plots the spatial autocorrelation value (Moran's I) for DIME, the Merriam's kangaroo rat, plotted as a function of neighborhood inverse weighted distance, and Figure 21 shows the same for surface temperature. In both cases, the spatial autocorrelation maintains a nearly constant, high plateau value at small distances, and drops off significantly at distances larger than 10,000 meters. Broad-scale physiography

of the study site matches this distance. For example, as one travels eastward through the study area, elevation modulates between the low values in the west-side Growler Valley, increases eastward across the Growler and Bates mountain ranges, back to a low elevation at the central Valley of the Ajo, and finally increases again to high elevation at the eastern Ajo Mountains. Dimensions of the each of the two valleys and two mountain ranges are approximately 10 km width. This broad-scale physiography is manifested in vegetation, as seen in the ReGap vegetation map (Figure 3), which also alternates between the large, approximately 10km expanses of the creosote-bursage class and the palo verde-mixed cacti class as one traverses across the same area.

In summary, broad-scale spatial autocorrelation is significant in both the unknown variable (rodent abundance) and in explanatory variables. Building a model with spatial corrections in either one alone is not justified. In fact, we suggest that one reason surface temperature is successful (or heavily used) in both the parametric model and non-parametric model is that it captures the broad-scale spatial trends at work in ORPI better than the other explanatory variables.

5. Discussion and conclusion

5.1 Comparison of SWR and CART methods

Surface temperature, or TSurf, was the variable with highest magnitude SWR coefficient (z-scored) in 4 out of the 10 species, and 8 out all 17 tests. The vegetation index, SAVI, was the second most expressed: 2 out 10 species and 5 out 17 tests.

Variables that operate at broad scales are partitioned first in CART; variables that act on fine scales appear near the terminal nodes. Surface temperature was partitioned

first in 4 of the 10 species, and 7 out of all 17 tests. Surface temperature was used to partition rodent plots in 11 of the total (38) binary splits in all ten species, more often than any of the other variable. The texture feature, Text_SAVI, a measure of pixel heterogeneity, and the vegetation index SAVI, were used 6 and 5 times, respectively.

Comparisons between SWR and CART results (Tables X and X1) in terms of selected predictor variables indicate a high level of uniformity. For example, 16 of the 38 variables for the ten species used in CART were also used in the corresponding species' SWR model. While SWR and CART differ in several ways, (SWR is a parametric regression test, while CART is a non-parametric, regressive method based on binary splits), similarities in variable selection indicate a convergence of models.

5.2 Surface temperature-based models

The frequent selection of surface temperature in both SW and CART, often the predictor with the highest coefficient magnitude in SWR and at the root node in CART, suggest it that it best characterizes the landscape of all the predictors. Its placement at the root node, and its spatial autocorrelation characteristics, suggest that is operating at a broad scale in the region.

We provide possible explanations of the success of surface temperature: 1) daytime surface temperature is a key factor in determining species habitat preferences or limits due to thermal tolerances; 2) similar tolerances of vegetation species on which rodent species depend for food or protection; and 3) it is an effective measure of a soil property, such as sand content, friability, or even amount of soil versus vegetation; and 4) it is an integrator of properties of several landscape components.

Due to the almost universal reliance upon surface temperature by all species and over all habitats in ORPI, we feel the evidence suggest the latter. Our results suggest that surface temperature is integrating effects of several biophysical factors, including soil, vegetation, slope, elevation and curvature, into a temperature function. Evidence of this is fairly high correlations between surface temperature and other variables at both 30-plot sampling and raster-wide comparisons (Figure 19).

Surface temperature correlated positively with curvature, creosote-bursage vegetation class, fine soils, sandy soils, soil erodibility factors, soil available water factors, and overall scene brightness. Overall scene brightness was expected to be negatively correlated with surface temperature (good reflectors are poor emitters of thermal energy, Kirchhoff's law or thermal radiation); however, an evaluation of scene components indicated terrain-generated effects such as shadowing altered this response. Surface temperature correlated negatively with slope, elevation, a clay soil spectral index, and rocky soils. Additional data sources for this table include the ReGAP vegetation image (ReGap, 2004), SSURGO soil table attributes (SSURGO, 2012), field data collected at the sites (Holm, 2006), and other spectral indices generated from the Landsat data set.

5.3 Overall model performances

A variety of methods are reported and available for habitat modeling (regression, regression tree, nonlinear, spatial, logistic, etc.), and modelers are presented with many choices. A first order choice may include the type of model. We chose stepwise regression and regression tree methods as suitable methods for the data due to their simplicity and straight-forward use and interpretation.

The methods and appear to be sufficient for the task as evidenced by convergence to simple, robust models with low error. SWR resulted in models with between two to four variables in all species. Likewise, CART produced models with between two and eight splits. Residual tree deviances ranged between 0.0065 and 0.0931 for the models, which are a fraction of the original deviance of the unsplit data set (1.00). SWR adjusted coefficient of determinations ranged between 0.43 and 0.83, with p-values ranging from 2.4E-3 to 2.6E-9. Given the constraints of a limited data set, these are small ranges considering such a large number and diverse set of species.

Second order choices include selection of default parameters (SWR entrance and exit parameters, and CART splitting tolerances, for example) and addition of interaction terms and spatial corrections. Manipulating default parameters affected the number of variables selected; however, by using criteria such as Akaike's Information Criterion (SWR) and cross-validation costs (CART), we felt the models converged to consistent, logical, and simple models. Addition of interaction terms, non-linear terms, and spatial autocorrelation corrections did not significantly improve the models.

5.4 Summary of predictions and comparison to literature

CHIN, or the rock pocket mouse, as its name implies, inhabits areas with rocky substrates or gravelly soil, and in or near rock ledges, fissures, piles, and the rocky slopes of mountains (Hoffmeister, 1986; Petryszyn and Cockrum, 1990; Flesch, 2008; Petryszyn, 1982; Lazaroff, 1998). Surface temperature was the dominant variable in both SWR and CART models, with either distance-to-wash or curvature attenuating. Areas with lower surface temperature value exhibited higher abundance. In separate

tests (Figure 19), surface temperature correlated negatively with increased slopes, and negatively with rocky, gravelly areas based on field and soil data.

PEAM (the Arizona pocket mouse) inhabits the sandy, open desert with sparse vegetation of grasses, mesquites, creosote bushes, and a few cacti (Hoffmeister, 1986; Petryszyn and Cockrum, 1990; Rosenzweig and Winakur, 1969; Lazaroff, 1998). It is abundant on valley floors and less as slope increases (Flesch, 2008). It is associated strongly with DIME in creosotebush areas with more or less level ground, and not common near mesquite thickets or large washes. It occurred in occurred in 15-35% perennial cover (least cover class) in one study (Stamp and Ohmart, 1978). SWR indicate that PEAM is found on westward facing slopes. Both SWR and CART models suggest areas with low vegetation content (SAVI), and in areas with higher surface temperature.

Literature suggests that DIME occupies low-lying, flat, open areas, with less vegetation (Stamp and Ohmart, 1978; Flesch, 2008; Petryszyn, 1982; Huggett, 1994). It does not occur on rocky slopes (Petryszyn and Cockrum, 1990; Flesch, 2008). It is observed in a variety of soil strengths and composition (Rosenzweig and Winakur, 1969; Flesch, 2008; Lazaroff, 1998; BLM, 2013) but may favor fine, sandy soils. The species is associated with the creosotebush (Monson and Kessler, 1940; Schroder, 1987). DISP, a larger kangaroo rat than DIME, prefers even more open, soil-exposed areas and less overall and shrubby vegetation (P, Hoffmeister, 1986; Rosenzweig and Winakur, 1969). Loose, sandy soil was not preferred as tunnels needed more support.

Both SWR and CART suggest DIME and DISP are found in areas with higher surface temperature. The DISP cutoff temperature in CART was even higher than

DIME. Temperature correlated negatively with elevation, slope, and positively with the ReGAP vegetation class of creosote-bursage. SWR also predicts that DIME is found in more southern and western facing slopes, or areas with more solar radiation and higher surface temperatures. A negative, concave curvature would be consistent with valleys. Higher abundance was predicted by CART in areas with low vegetation (SAVI < -0.362) and lower texture or pixel diversity, which is consistent with homogenous expanses of creosote flats.

NEAL, the wood or pack rat, and the largest murid species, is found in a variety of habitats. Literature suggests that NEAL is associated with areas of high cacti such as prickly pear and cholla, which is used for food, a water source, and building material for ground surface level dens (Hoffmeister, 1986; Petryszyn and Cockrum, 1990; Petryszyn, 1982; Lazaroff, 1998; Monson and Kessler, 1940). Areas with mesquite are often favored due to ample food from mesquite pods. NEAL is associated with areas of higher slope, and not in areas of pure creosote stands (Flesch, 2008; Lazaroff, 1998). PEER, the cactus mouse, is generally observed in similar habitats, and especially in rocky, sloped areas.. It is the most common rodent in desert areas with denser vegetation (Petryszyn and Cockrum, 1990; Hardy, 1974). In contrast to DIME, it cannot run fast, but is good at climbing and scrambling, and for this reason it lives in rocky terrain or sometimes in areas with cactus or brush, where safe havens are within clambering distance (Flesch, 2008; N; Rosenzweig and Winakur, 1969).

SWR resulted in almost identical predictions for NEAL and PEER. Both display similar coefficients and a negative correlation with surface temperature, and a positive correlation with texture (pixel or landcover diversity). CART suggests higher NEAL

abundance at higher sloped areas (node 1, broad-scale differentiation), and in higher textured portions of areas with lower slope. Higher PEER abundance is suggested by CART in lower surface temperature and areas with high pixel texture.

Agreement between the models and literature for NEAL and PEER is suggested by negative correlations of surface temperature with vegetation content (SAVI) and with the mixed-cacti paloverde ReGAP class, and positive correlation with sloped areas. Higher abundance is indicated also in areas with diversity (texture) by both models, in agreement with literature.

ONTO, the southern grasshopper mouse, is the southern species of grasshopper mice found in Arizona (Hoffmeister, 1986). The species is noted for preferring patchy vegetation and more low-lying, xero-riparian areas with exposed soil (Hoffmeister, 1986; Flesch, 2008; Hardy, 1974). UTM northing, an equivalent to latitude, was used in both SWR and CART. Negative correlation and higher abundance with more southern coordinates suggest a preference by ONTO for southern areas of the study area, consistent with its name. Only the ONTO models called up a UTM coordinate, northing in this case, as the prominent variable. SWR also suggest higher abundance in areas with high texture or landscape diversity.

SWR did not generate a statistically significant model for SIAR, the cotton rat, as it was captured in too few of the rodent plots. Literature suggests SIAR is found in diverse habitat, and often along waterways, drainages and agriculture, and in areas with stands of grass (Hoffmeister, 1986; Lazaroff, 1998; Jorgensen et al., 1998). CART model suggests higher abundance is found on southwest slopes in the study area, especially areas with higher surface temperature or areas with higher vegetation.

5.5 Advantages and limitations of the methods

The predictor variable set for this study was derived from readily available, free data sources. Image processing for most of the spectral and topographical data was minimal. The surface temperature data necessitated substantial processing time. However, pixel-by-pixel comparison in the study area between the surface temperature image and the original Landsat band 6 image from which it was derived indicate very high correlation (+0.97), suggesting that the additional processing was not necessary.

Spatial resolution is limited to 30 m pixels using Landsat images. Finer scale habitat characteristics, such as a rodent burrow complex, a group of creosote bushes, an individual mesquite tree or boulder, will not be discernible. However, the pixel size is well suited for the dimensions of the rodent plots. In fact, we found better results using 5x5 median low pass filter to average neighborhood pixel values. The pixel size is also well-suited for data management of a regional study's extent, such as that of ORPI.

The models developed in this study are static temporarily, and reflect an average or overall abundance and biophysical conditions between the years captured by the data sets (1993-2006). They do not reflect ranges or diversity of the data. However, models designed to compare inter-annual differences and long term changes are possible with the methods and data described here.

Predictions based on these models are limited to the range of the individual predictor variables measured across the 30 plots. Extrapolation beyond this range would require further testing. Predictions based of the models are also limited to the area studied, western Pima County, AZ in the Sonoran Desert. For example, a model for kangaroo rats built with ORPI surface temperature, aspect and curvature would not be

appropriate for the Mojave Desert due to different landscape characteristics and their spectral responses. However, the model shows promise for use in analogous regions in the Lower Sonoran Life Zone with similar physio-geography, such as eastern Pima County.

5.6 Potential applications of the models

Coccidioidomycosis, or Valley Fever, is endemic in regions of the Lower Sonoran Life Zone. Desert rodents have long been suspected as playing a role in coccidioidomycosis transmission via numerous pathways (Tabor, 2009). A first step to shed light on this hypothesis is to determine where rodents live.

We expect that models described here will allow us to generate rodent abundance maps based on value or on ranked classes for the ORPI study area. We will test and extrapolate to nearby, analogous regions in the Lower Sonoran Life Zone within the range of the predictor variable set. Comparisons of exogenous variable ranges between ORPI and the surrounding area suggest that the models calibrated to ORPI are representative of the regional biome. We further aim to use rodent abundance data, associated with human disturbance metrics derived from dust generating construction, to analyze potential rodent contributions to coccidioidomycosis incidence in Pima County, Arizona.

6. Acknowledgments

We would like to express major appreciation to the Organ Pipe Cactus National Monument, and in particular, Peter Holm, for lending us their data set and providing input on their rodent monitoring methods. We also would like to thank Dr. Daoquin

Tong at the Department of Geography and Development, University of Arizona for her assistance with statistical aspects of the models, and to Dr. Stuart Marsh at the School of Natural Resources and the Environment, University of Arizona for his assistance with remote sensing evaluation.

7. Figures and tables

Figure 1: The Organ Pipe Cactus National Monument is situated near the geographical center of the Sonoran Desert (striped area). It lies in western Pima County, Arizona, and it shares its southern boundary with the border with Mexico. It lies approximately 240 km east of the city of Tucson (yellow). The spatial spread of the rodent plots sites is shown as points.

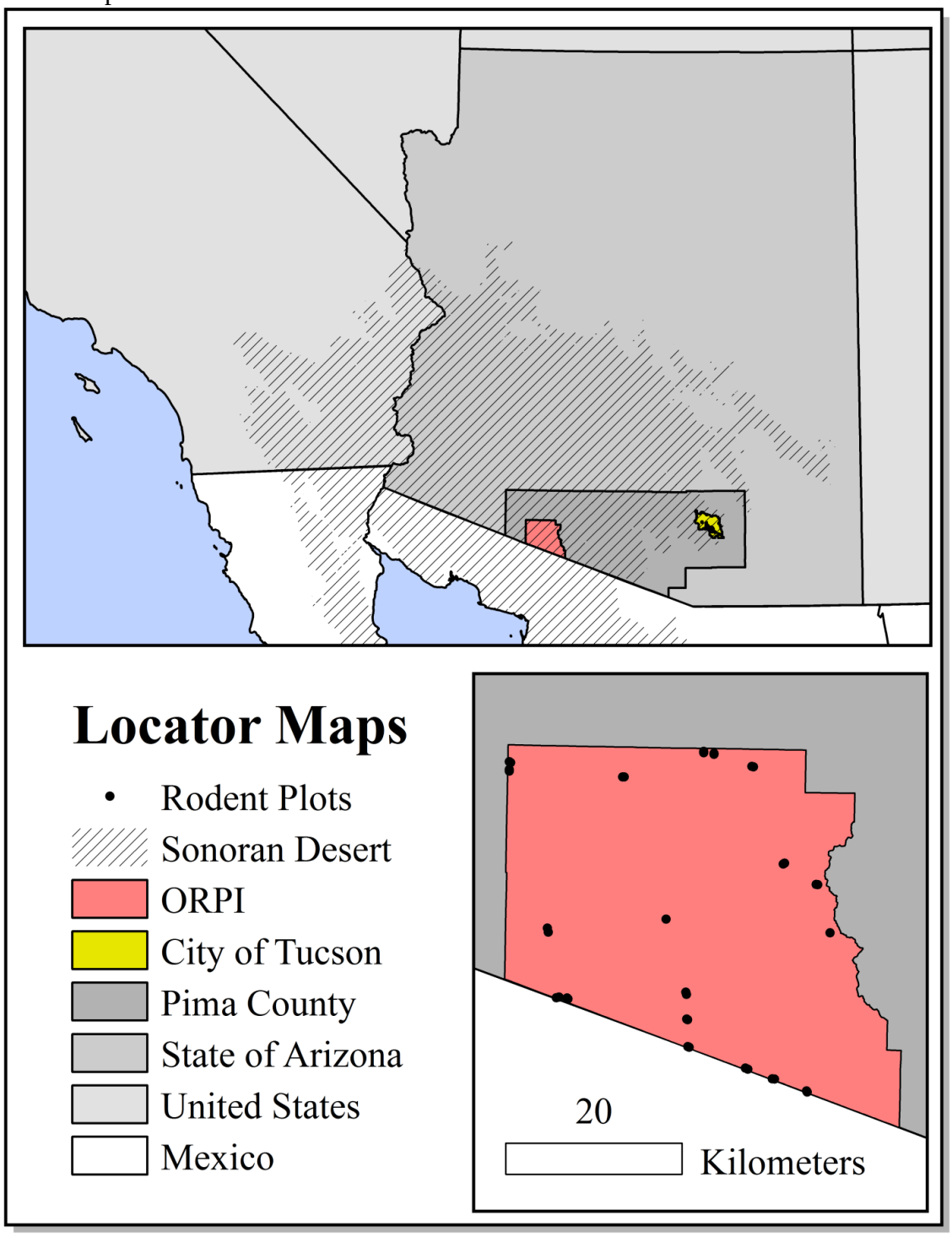


Figure 2. Political map of Southwest Arizona showing major land administrational areas. The ORPI is located in a sparsely-inhabited region. Arizona State Highway 85 bisects the monument and links the state with the Gulf of California in Mexico. Sonoyta, Sonora (Mexico) and Ajo, Arizona are the largest nearby cities.

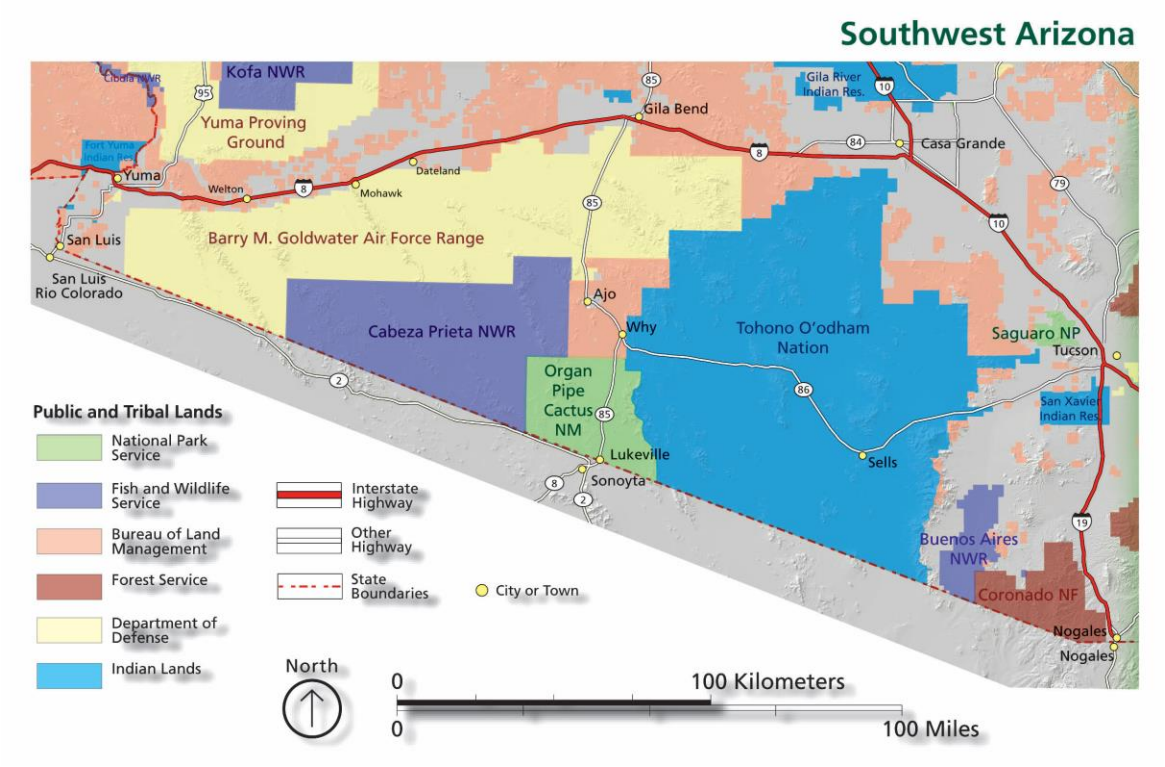


Figure 3: Vegetation classes derived from Southwest Regional GAP data set. Landcover classes in ORPI derived from the Southwest Regional GAP data set (ReGAP, 2006). The ReGAP landcover maps are produced from Landsat ETM+ imagery in conjunction with a digital elevation model (DEM). ReGAP landcover classes are drawn from NatureServe's Ecological System concept, decision tree classifiers, and other techniques. The ORPI consists of broad expanses of two classes: Sonoran-Mojave creosotebush-white bursage desert scrub (brown or dark) and Sonoran palo verde-mixed cacti desert scrub (light green or light). Significant agriculture use is located south of ORPI's border with Mexico, and upland chaparral and woodland are present in the higher elevations of the Ajo Mountains on the east side.

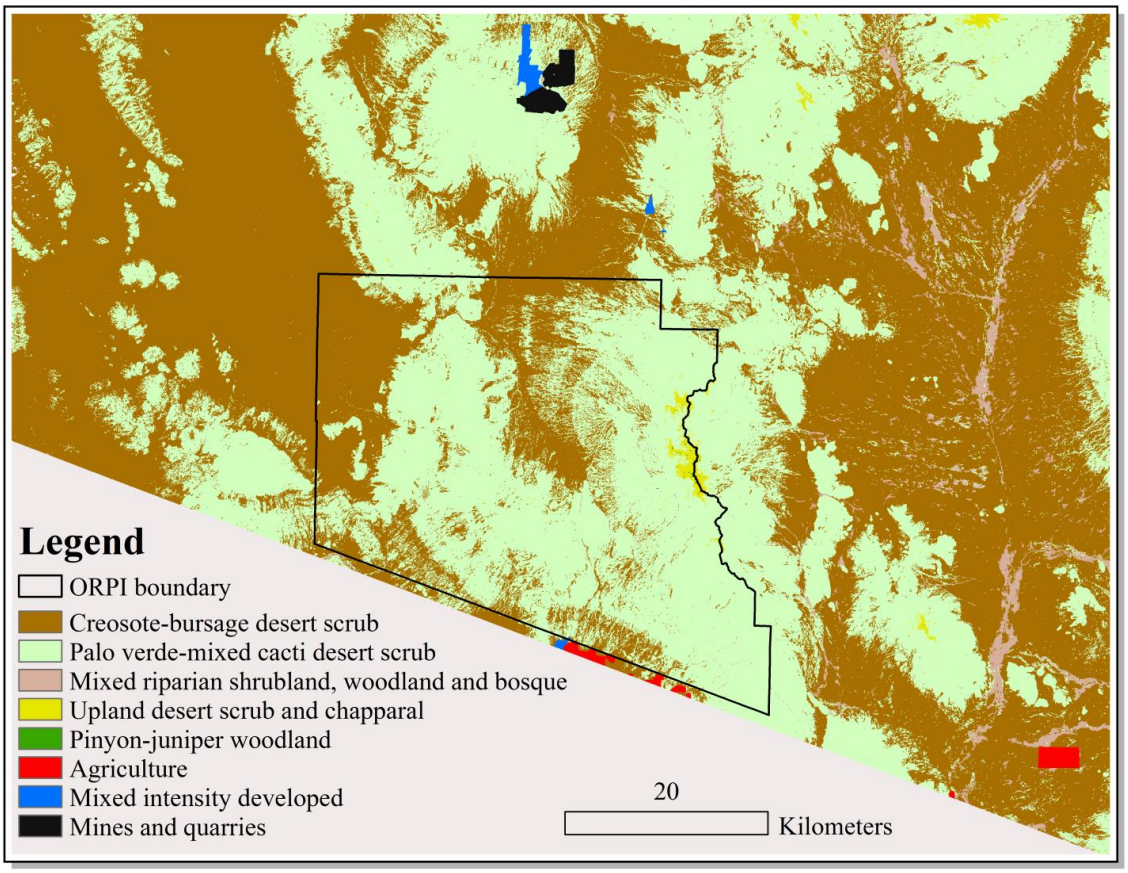


Figure 4: Summary of rodent abundance data for individual species and groupings. Abundance data were compiled from Organ Pipe Cactus National Monument data obtained by information request, and represents fourteen year means of annual capture and release trapping events from 1993-2006.

Code	Species	Common Name	Abundance Range	Abundance Mean	Abundance Standard Deviation
CHBA	*Chaetodipus baileyi*	Bailey's pocket mouse	0.1-34.3	6.6	10.7
CHIN	*Chaetodipus intermedius*	rock pocket mouse	0.0-16.4	2.1	4.5
CHPE	*Chaetodipus penicillatus*	desert pocket mouse	1.4-70.9	21.5	18.9
PEAM	*Perognathus amplus*	Arizona pocket mouse	0.0-19.9	5.8	5.4
DIME	*Dipodomys merriami*	Merriam's kangaroo rat	0.0-27.9	12.5	9.3
DISP	*Dipodomys spectabilis*	banner-tail kangaroo rat	0.0-0.6	0.1	0.2
NEAL	*Neotoma albigula*	white-throated woodrat	0.0-20.2	4.1	5.1
ONTO	*Onychomys torridus*	grasshopper mouse	0.0-1.8	0.3	0.4
PEER	*Peromyscus eremicus*	cactus mouse	0.0-12.0	1.3	2.8
SIAR	*Sigmodon arizonae*	Arizona cotton rat	0.0-0.8	0.0	0.2
Het	Heteromyids	CHBA, CHIN, CHPE, PEAM, DIME, DISP	25.7-85.9	48.6	15.7
Mur	Murids	NEAL, PEER, SIAR, ONTO	0.0-32.2	5.8	7.6
Total	All Species		28.9-88.6	54.4	15.1
Val	Valley types	CHPE, PEAM, DIME, DISP, ONTO	2.6-86.2	40.2	23.6
Roc	Rocky types	CHBA, CHIN, NEAL, PEER, SIAR	0.1-57.9	14.2	18.2
Total BM	Total Biomass*		40.1-112.8	74.3	19.6
SDI	Shannon's Diversity Index*	$H' = -\Sigma(p_i * \ln(p_i))$	0.6-1.3	1.0	0.2
n	Richness	Number of species	2.8-5.3	4.0	0.6

Figure 5: ORPI rodent monitoring plot dimensions. A comparison of the dimensions of two neighboring rodent plots in ORPI. The Lower Colorado Larrea (LOWE) plots 1 and 2, located in northwest ORPI, are shown in black outline with a black dot at their respective centers. The plots are overlain the 14-year median Soil Adjusted Vegetation Index image (SAVI), whose pixel dimensions are 30m. Higher SAVI values are represented by increased lighter shades of gray. The dimensions of the low pass spatial filter window (5x5 pixel median value) used to average the SAVI are shown in dotted red for comparison with plot dimensions (approximately 2.25 ha and 0.8 ha filter respectively). The predictor variable value for each plot is sampled from the spatially-filtered image at the center pixel of the plot.

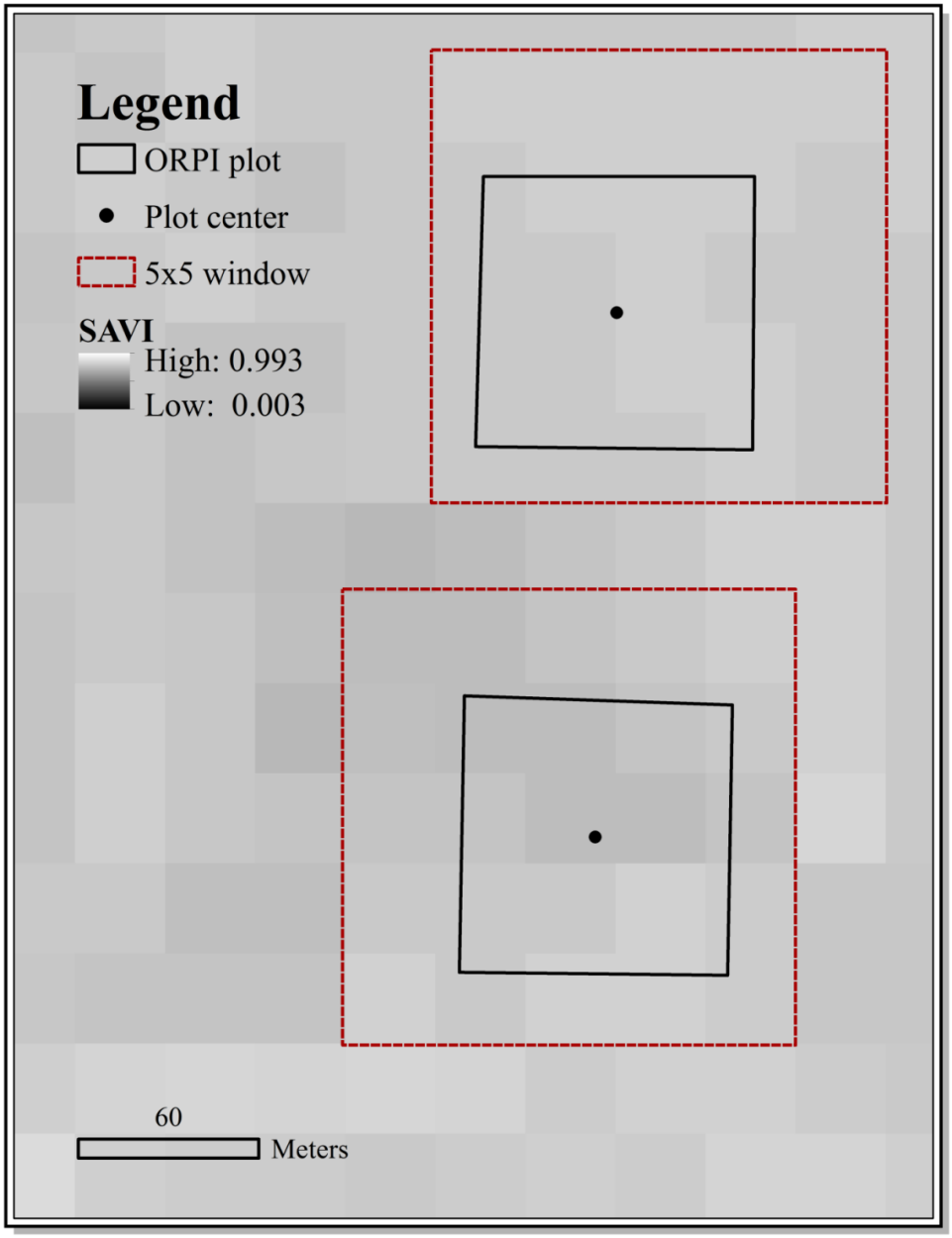


Figure 6: Predictor variable ranges. The biophysical data shown is derived from Landsat TM satellite imagery and digital elevation models, and it is used to build predictive models for rodent abundance. Statistics are determined from 30-point sampling (rodent plots) of 5x5 median-filtered images.

Code	Biophysical Variable	Range Over 30 plots	Mean	Standard Deviation
TSurf	Surface Temperature	318.2 - 332.3 (K)	326.73	2.93
SAVI	Soil Adjusted Vegetation Index	0.582 - 0.632	0.601	0.0132
Text_SAVI	Texture SAVI	1.60E-6 - 3.42E-4	4.54E-05	8.45E-05
UTM_E	Easting Coordinate	303548 - 339927 (m)	323516	11528
UTM_N	Northing Coordinate	3525190 - 3563822 (m)	3545986	13753
Elev	Elevation	327 - 954 (m)	482	146
Slope	Slope	0.17 - 22.70 (°)	3.24	5.02
Asp_E	Aspect Easting	-0.994 - +0.966	-0.584	0.602
Asp_N	Aspect Northing	-0.940 - +0.682	-0.148	0.418
Curv	Curvature	-0.532 - +0.112	-0.0217	0.1075
Dist_Wash	Distance to Wash	0 - 268 (m)	84.2	63.2

Figure 7: Histogram showing the frequency distribution of Merriam's kangaroo rat (DIME) abundance at 30 rodent plots in ORPI. Abundance of DIME ranges from 0 to 27.9 at thirty rodent trapping sites. Abundance values are evenly distributed over the range.

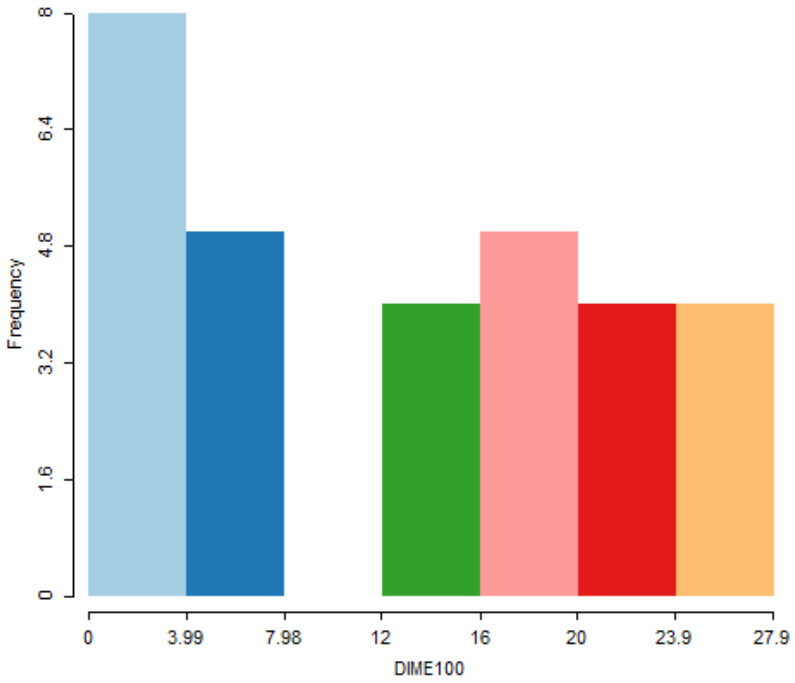


Figure 8: Histogram showing the frequency of surface temperature values. Surface temperature is derived from Landsat Band 6 (thermal infrared), and sampled at the 30 rodent plots in ORPI. Values are standardized. Z-scored surface temperature ranges from -2.9 to 1.91, or 318.2 - 332.3 (K). Values are roughly normally distributed, but with a right-leaning skew.

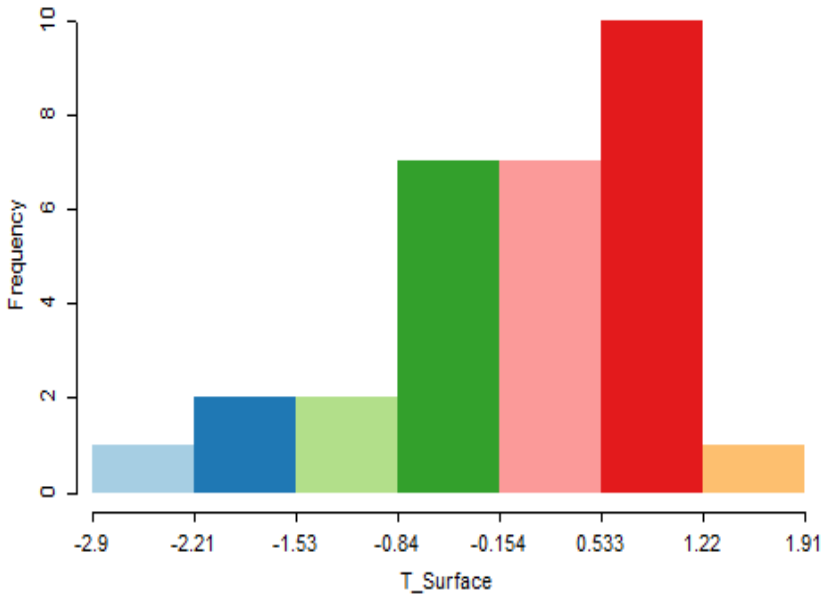


Figure 9: Scatterplot comparing the relationship between Merriam's kangaroo rat (DIME) abundance and surface temperature derived from Landsat thermal infrared imagery. The values are sampled from 30 rodent plots located in the Organ Pipe Cactus National Monument. A statistically significant linear relationship is shown (adjusted coefficient of determination of 0.67).

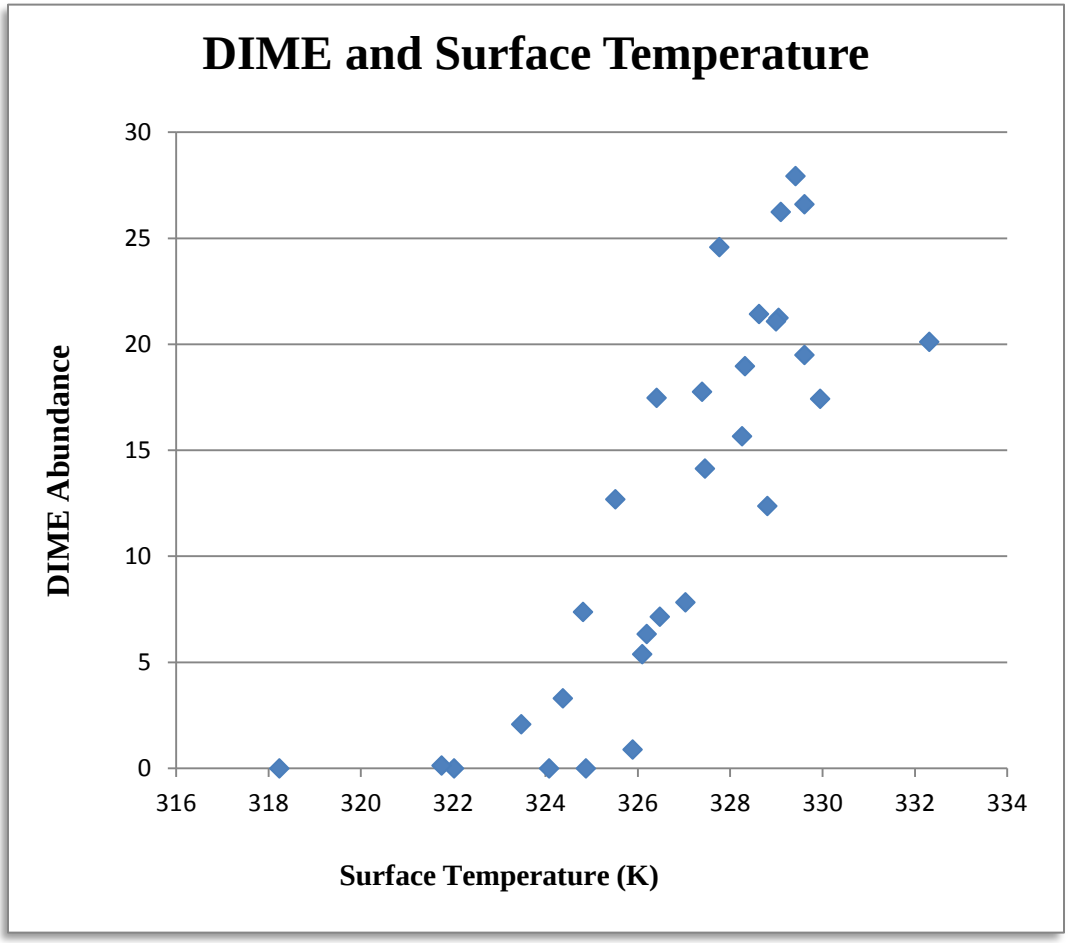


Figure 10: Model building. Adjusted coefficient of determination values for various combinations of predictor variables. For example, using only surface temperature as an explanatory variable resulted in statistically significant (p-value < 0.05) models for only three species, CHIN, DIME, and NEAL. A model composed of the suite of topographical variables (elevation, slope, aspect and curvature) plus latitude and longitude results in only models for only two of the species. The addition of the surface temperature variable to this set results in six more species arriving at a statistically-significant endpoint. We added SAVI and texture SAVI for our final set of variables. The model shown in the far right column adds additional spectral indices and resulted in improved adjusted R^2 values; however we did not use these due to an assessment of AICc comparisons and over-fitting of the data.

	tsurf	**savi**	**tsurf +savi +textsavi**	**topo +utm**	**topo +utm +tsurf**	**topo +utm +tsurf +savi +texsavi**	**topo +utm +tsurf +savi +texsavo +3spectral**
Species	**Adj R2**	**Adj R2**	**Adj R2**	**Adj R2**	**Adj R2**	**Adj R2**	**Adj R2**
Chba	-	-	-	0.63	0.63	0.69	0.77
Chin	0.56	-	0.56	-	0.56	0.56	0.79
Chpe	-	-	-	0.48	0.48	0.56	0.65
Peam	-	-	-	-	0.42	0.42	0.58
Dime	0.66	-	0.66	-	0.80	0.80	0.81
Disp	-	-	-	-	-	-	-
Neal	0.51	-	0.63	-	0.56	0.63	0.77
Onto	-	-	-	-	-	-	-
Peer			0.74	-	-	0.74	0.88
Siar	-	-	-	-	-	-	-
Het	-	-	0.44	-	-	0.44	0.60
Mur	-	-	0.69	-	0.55	0.69	0.82

Figure 11: Abundance model predicted by stepwise linear regression for CHBA (Bailey's pocket mouse). The overall adjusted coefficient of determination is 0.67 (p-value 1.8(E-6)). Elevation is predicted to have the highest impact, followed by aspect (northing) and distance from wash. All have positive effect.

Variable	**Coefficient**	**Standard Error**	**p-value**
Elevation	0.567	0.141	0.0004
Aspect_N	0.376	0.135	0.0097
Dist_Wash	0.272	0.123	0.0353
Constant	0.000	0.110	1.0000

Figure 12: Stepwise multiple regression summary of rodent abundance models with adjusted coefficient of determinations, associated p-values, predictors and their coefficients for rodent species. Negative coefficients are highlighted. Note: CHIN, DISP, ONTO, PEER, SIAR data from reduced plot models.

Species/Group	Adj R Sq	p-value	intercept	Coefficient 1	Predictor 1	Coefficient 2	Predictor 2	Coefficient 3	Predictor 3	Coefficient 4	Predictor 4
CHBA	0.67	1.77E-06	-6.70E-17	0.57	Elevation	0.38	Aspect_N	0.27	Dist_Wash		
CHIN	0.75	1.85E-04	-5.00E-15	-0.86	T_Surface	0.39	Dist_Wash				
CHPE	0.53	1.85E-04	8.00E-15	0.49	SAVI	-0.47	Aspect_N	0.35	T_Surface		
PEAM	0.66	2.48E-06	7.04E-15	-0.55	SAVI	0.39	UTM_N	0.34	T_Surface		
DIME	0.83	2.58E-09	1.83E-14	0.94	T_Surface	-0.28	Aspect_E	-0.22	Aspect_N	-0.22	Curvature
DISP	-	-	-		-						
NEAL	0.65	6.16E-07	-1.07E-14	-0.55	T_Surface	0.40	Text_SAVI				
ONTO	0.43	2.421E-03	1.11E-14	-0.66	UTM_N	-0.25	Aspect_N				
PEER	0.74	1.584E-05	-8.90E-15	-0.53	T_Surface	0.50	Text_SAVI				
SIAR	-	-	-	-	-						
Heteromyids	0.48	1.56E-04	1.31E-14	0.61	T_Surface	0.53	SAVI				
Murids	0.71	5.09E-08	-1.00E-14	-0.52	T_Surface	0.48	Text_SAVI				
All Species	0.29	2.20E-03	1.34E-15	0.54	SAVI						
DIME + PEAM	0.78	1.22E-09	1.61E-14	0.83	T_Surface	-0.38	Aspect_E				
Valley types	0.73	1.66E-07	1.40E-14	0.69	T_Surface	-0.47	Aspect_N	0.25	SAVI		
Rocky types	0.82	7.94E-10	-1.44E-14	-0.74	T_Surface	0.42	Aspect_N	0.22	Aspect_E		
Heteromyid BM	0.62	1.31E-05	1.25E-14	0.67	SAVI	0.56	T_Surface	-0.35	Text_SAVI		
Murid BM	0.70	1.02E-07	-1.03E-14	-0.53	T_Surface	0.46	Text_SAVI				
Total BM	0.32	1.10E-03	1.41E-15	0.57	SAVI						
DIME + PEAM BM	0.79	6.51E-10	1.63E-14	0.84	T_Surface	-0.36	Aspect_E				
Valley Types Biomass	0.76	3.23E-08	1.46E-14	0.72	T_Surface	-0.47	Aspect_N	0.24	SAVI		
Rocky Types Biomass	0.87	1.14E-10	-1.47E-14	-0.70	T_Surface	0.34	Dist_Wash	0.32	Aspect_N	0.30	UTM_E
SDI	0.61	6.14E-05	-1.32E-14	-0.81	SAVI	0.49	UTM_E	-0.48	T_Surface	0.43	Text_SAVI
Exponent SDI	0.63	3.82E-05	-1.39E-14	-0.77	SAVI	-0.52	T_Surface	0.49	UTM_E	0.41	Text_SAVI
Number Species	0.31	1.40E-03	-1.08E-14	-0.56	T_Surface						

Figure 13: Regression tree for CHBA. Statistics for five terminal nodes are shown in the boxes. The root node or first split calls up aspect-northing. Those observation (rodent plots) with aspect northing of z-score less than 1.14637 (+0.3311 actual aspect value) are passed to the left, and those equal greater are passed to the right. Both sides of the second-level splitting employ surface temperature. Texture of the SAVI image (diversity) is used as the final third level split. Terminal nodes are nodes 5-9.

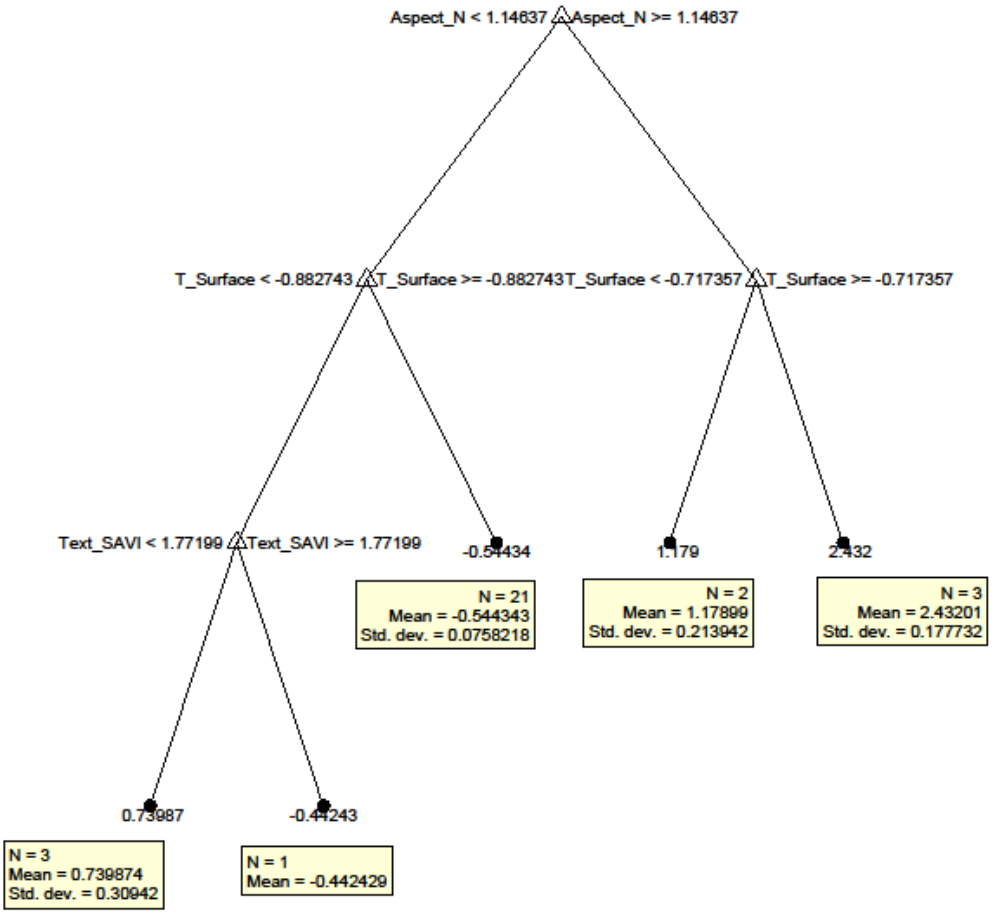


Figure 14: Summary of regression tree (CART) model for the CHBA (Bailey's pocket mouse). Statistics for each terminal node are shown. The number of cases is the number of rodent plots out of 30 total represented by this grouping. Deviance is a measure of heterogeneity in CHBA abundance amongst the rodent plots grouped into the terminal node. The predicted values are standardized (z-scored) abundance for CHBA. Node 7 showing the highest abundance, and it represents areas with aspect northing greater than 1.15 (z-score) and surface temperature greater than -0.72 (z-score).

Node Number	Number of Cases	Deviance (Residual sum of Squares)	Predicted Value
5	21	0.1150	-0.544
6	2	0.0458	1.179
7	3	0.0632	2.432
8	3	0.1915	0.740
9	1	0.0000	-0.442

Figure 15: Variable importance in the Bailey's pocket mouse (CHBA) regression tree. Variable importance is the estimates of input feature importance for tree t by summing changes in the risk (node error weighted by node probability) due to splits on every feature. Variable importance associated with each split is computed as the difference between the risk for the parent node and the total risk for the two children. Aspect-northing is the dominant variable, followed by surface temperature, for the CHBA regression tree.

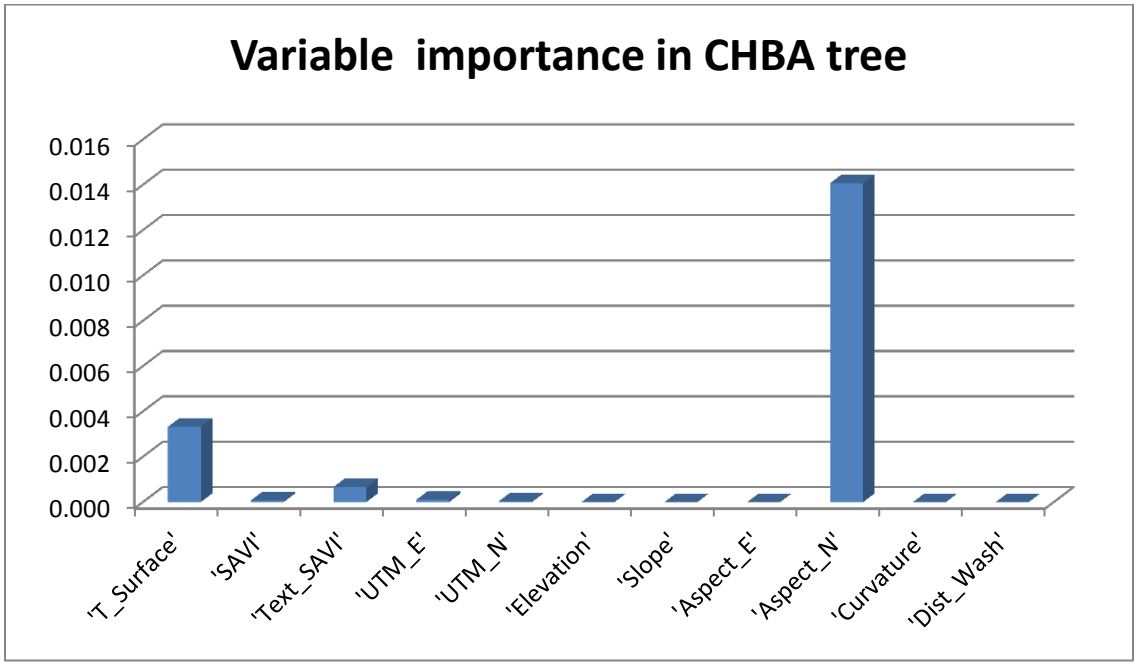


Figure 16: Overall tree deviance as a function of number of terminal nodes for the Bailey's pocket muse (CHBA). Tree deviance is calculated as the sum residual squares (sum of squares of observed – mean in the terminal nodes) divided by the degrees of freedom (number of total observations – number of terminal nodes). Note that the deviance drops drastically from a value of 1.0 in the root node (no splits) to approaching zero at five terminal nodes. By deviance criteria alone, the best size tree is pruned to five terminal nodes.

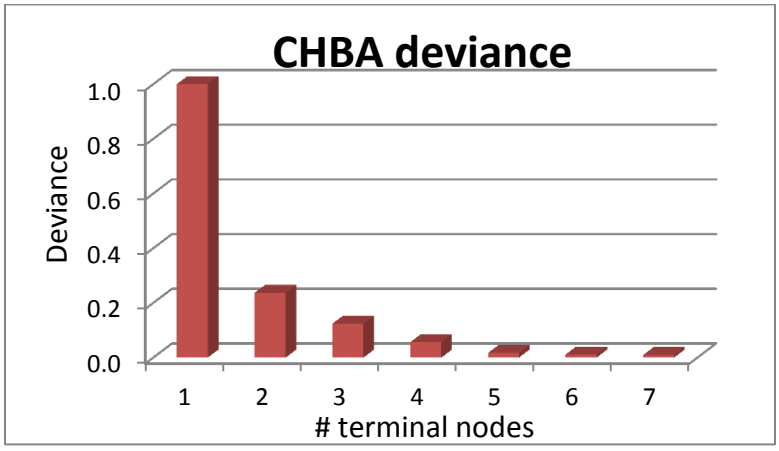


Figure 17: CHBA cross-validation cost analysis. The function uses 10-fold cross-validation to compute the cost vector. The cost of the tree is the sum over all terminal nodes of the estimated probability of a node times the cost of a node. The function partitions the sample into 10 subsamples, chosen randomly but with roughly equal size. For each subsample, test fits a tree to the remaining data and uses it to predict the subsample. It pools the information from all subsamples to compute the cost for the whole sample. This is one way to perform validation of model results, particularly when no additional data sets are available for comparison. In the CHBA case shown here, cross-validation cost testing suggests that a tree pruned to four or five terminal nodes is best suited to minimize overall deviance, and larger tree sizes do not improve statistics adjusted for tree size.

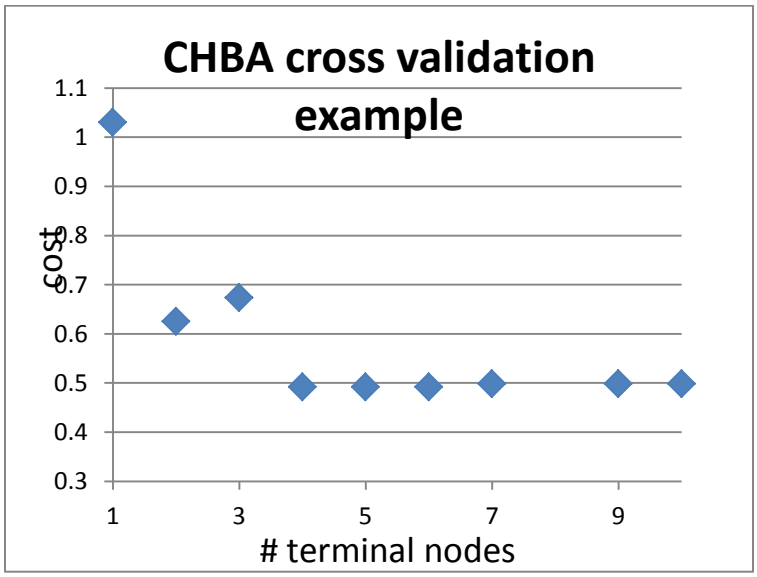


Figure 18: Regression tree summary for ten rodent species. Models are generated with a set of biophysical variables including surface temperature (TSurf), a vegetation index (SAVI), texture or pixel heterogeneity (Text_SAVI), topographical variables generated from a digital elevation model (elevation, slope, aspect, curvature, and distance from wash), and geographical coordinates (UTM_E and UTM_N). The residual deviance and resubstitution costs are measures of the unwanted heterogeneity represented in the terminal nodes of the overall tree. The main predictor is the first split in the tree, or root node, and other predictors are subsequent nodes used to divide and group abundance values.

Rodent Model	Terminal Nodes	Residual deviance (tree)	Resubstitution Costs	Main predictor	Other predictors
CHBA	5	0.0166	0.0138	Asp_N	TSurf, Text_SAVI
CHIN	4	0.0365	0.0316	TSurf	Curv, TSurf
CHPE	9	0.0754	0.0528	Text_SAVI	UTM_N, TSurf, SAVI, UTM_E
PEAM	8	0.0931	0.0683	Asp_E	SAVI, TSurf
DIME	7	0.0437	0.0335	TSurf	UTM_N, Text_SAVI, SAVI, ASP_N, TSurf
DISP	6	0.0289	0.0231	TSurf	SAVI, UTM_E, Elev
NEAL	6	0.0574	0.0459	Slope	Text_SAVI, UTM_E, Elev, TSurf
ONTO	5	0.0544	0.0454	UTM_N	Curve, Elev, Text_SAVI
PEER	5	0.0065	0.0054	TSurf	Text_SAVI
SIAR	4	0.0213	0.0185	Asp_N	TSurf, SAVI

Figure 19: Surface temperature cross-correlations. Comparison of surface temperature values with other biophysical variables sampled at 30 rodent plots, and comparisons between the pixels of entire images. Values are correlation coefficients with either a negative or positive relationship. Shown are correlations between surface temperature and other variables used to build models, and correlations with other data sources including field data, vegetation maps, soil maps and their attributes, and spectral indices. Correlation comparisons are useful in interpreting complexities between variable relationships and the sources of a variable's predictive power.

Variable	30 rodent plot sample	ORPI area images
TSurf	1.00	1.00
Unprocessed Landsat Band 6 TIR	0.98	0.97
SAVI	-0.27	0.02
Text_SAVI	-0.44	-0.10
UTM_E	-0.06	
UTM_N	0.13	
Elev	-0.48	-0.21
Slope	-0.73	-0.68
Asp_E	0.07	-0.03
Asp_N	-0.07	-0.03
Curv	0.54	-0.02
Dist_Wash	-0.06	-0.28
Canopy Cover	-0.58	-0.27*
ReGAP Creosote-Bursage Class	0.41	0.48*
Tasselled-Cap Brightness	0.59	0.28
Clay Minerals (LS B5/B7)	-0.59	-0.48
Iron Oxide (LS B3/B1)	0.34	0.44
Ferrous Minerals (LS B5/B4)	-0.13	0.05
T Factor[a]	0.82	
15 bar[a]	0.82	
Kw factor[a]	0.80	
third bar[a]	0.77	
restriction[a]	-0.65	
AWC[a]	0.73	
sandfine[a]	0.72	
topo valleyfloor[b]	0.75	
soil fine[b]	0.75	
RVSand[a]	0.74	
RVSilt[a]	-0.84	
RVClay[a]	0.11	
aridsols[a]	-0.79	
albedo dry[a]	-0.32	
wind erodibility[a]	0.89	
soil coarse[b]	-0.75	
rocky slope[b]	-0.77	
high range slope[a]	-0.80	

[a] SSURGO (2012) soil table attributes, top horizon layer.

[b] Field observations at rodent plot (Holm, 2006). Topography classes: valley floor, bajada and rocky slope; soil classes: fine and coarse.

Figure 20: Spatial autocorrelation as a function of neighborhood distance: a neighborhood distances of up to 10km, the Global Moran's I, a measure of spatial autocorrelation, for the Merriam's Kangaroo Rat (DIME) is statistically significant, and positive, indicating spatial clustering of abundance values. The spatial autocorrelation drops to almost zero at distances between 10 and 25km. The pattern is informative on the response of DIME's response to underlying spatial structure in the terrain.

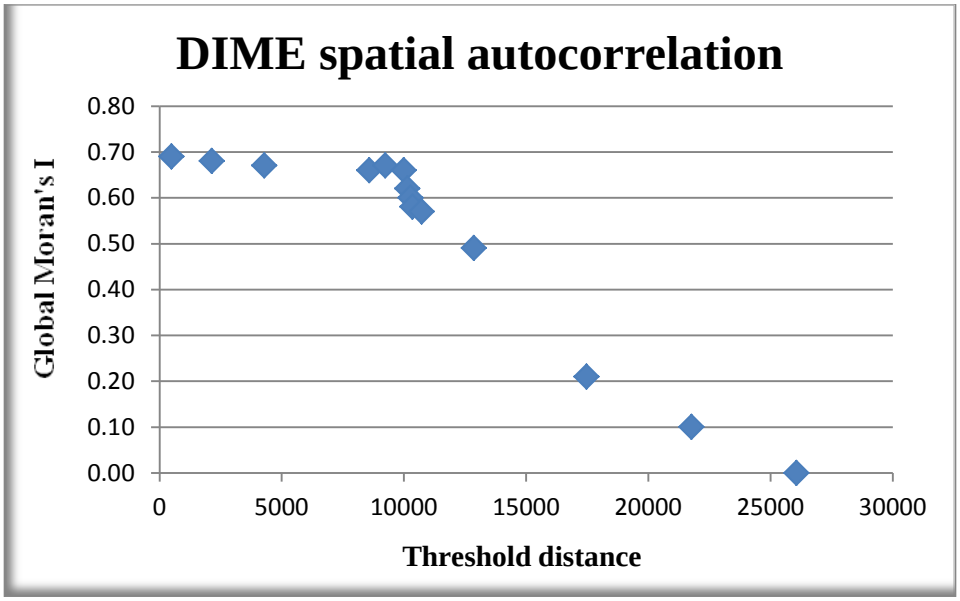
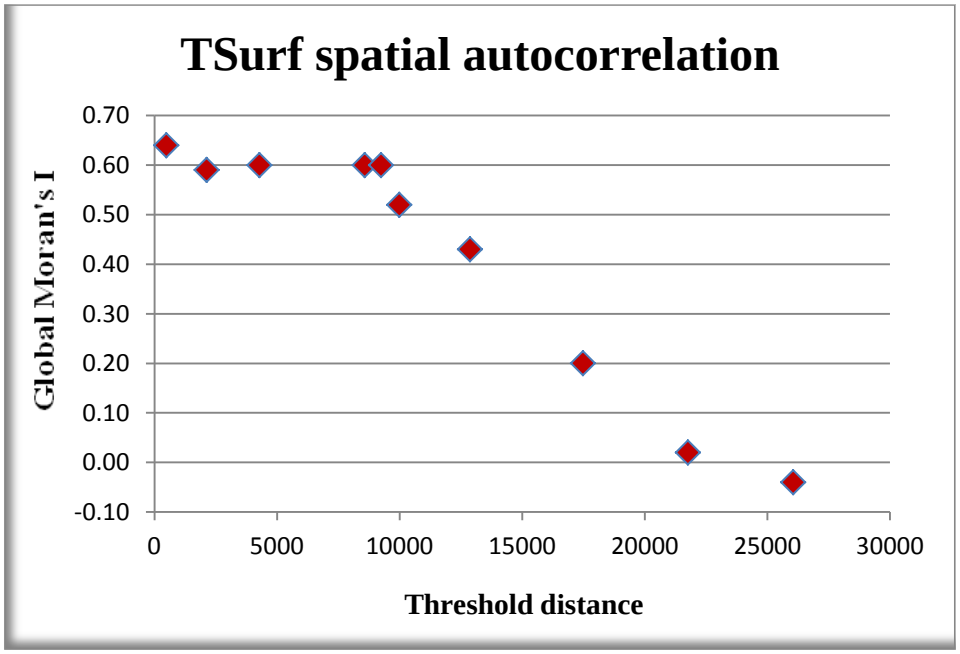


Figure 21: Spatial autocorrelation as a function of neighborhood distance: surface temperature exhibits a strikingly similar pattern in spatial autocorrelation as the kangaroo rat (DIME, Figure X-1). The similarity is informative on the reasons for the ability of surface temperature to predict DIME abundance.



8. Appendices

Appendix 1: Equations for determination of the surface temperature image.

1. $NDVI = (\rho_{nir} - \rho_{red})/ (\rho_{nir} + \rho_{red})$
where ρ is the respective spectral band reflectance

2. $FVC = (NDVI - NDVI_s) / (NDVI_v - NDVI_s)$
where
NDVI refers to the Normalized Difference Vegetation Index
$NDVI_s$ refers to the soil NDVI in the region (0.05)
$NDVI_v$ refers to the vegetation NDVI in the region (0.75)

3. $\varepsilon = \varepsilon_s*(1-FVC) + \varepsilon_v*(FVC)$
where
FVC refers to Fractional Vegetation Component
ε_s refers to emissivity of soil in the region (0.958)
ε_v refers to emissivity of vegetation in the region (0.952)

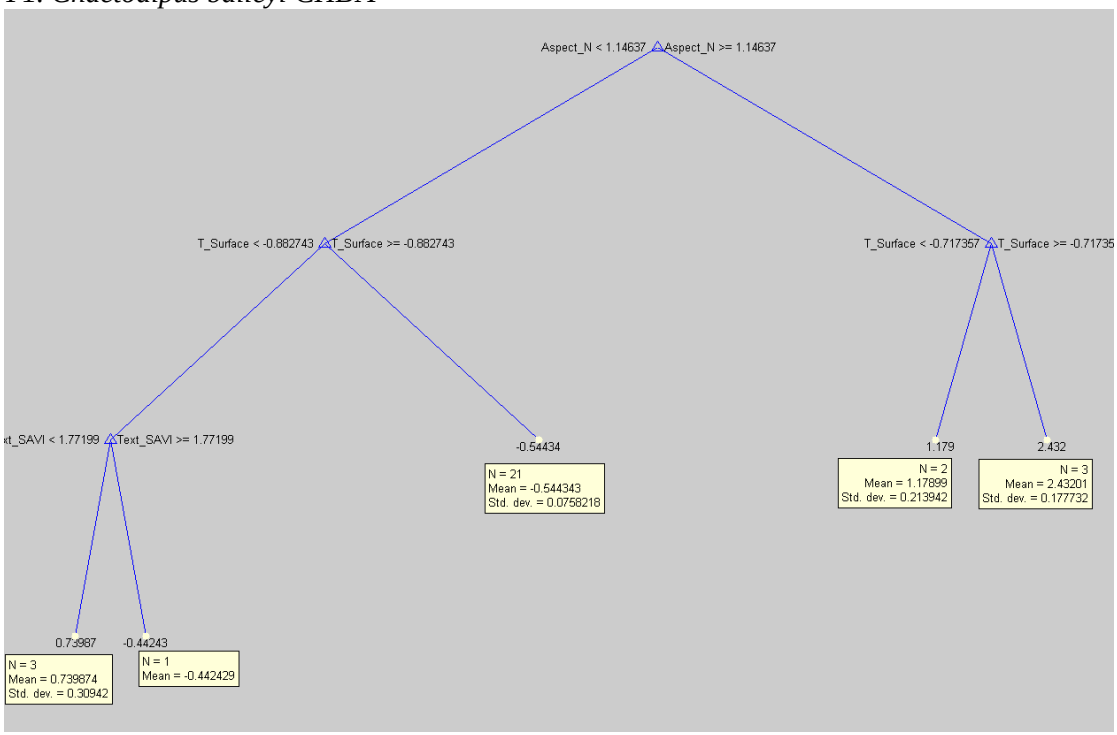
4. $L_\lambda = G_{rescale} \times Q_{cal} + B_{rescale}$
where
L_λ refers to the at-sensor top of atmosphere spectral radiance (W/m^2 sr μm)
Q_{cal} refers to the quantized calibrated pixel value (DN)
$G_{rescale}$ refers to the band specific rescaling gain factor (0.055376 W/m^2 sr μm/DN)
$B_{rescale}$ refers to the band specific rescaling bias factor (1.18 W/m^2 sr μm)

5. $LT_\lambda = (L_\lambda - L_u - \tau*(1-\varepsilon)*L_d) / (\varepsilon*\tau)$
where
LT_λ refers to the blackbody spectral radiance of surface
L_λ refers to at-sensor, top of atmosphere spectral
L_u refers to upwelling, atmospheric path radiance
L_d refers to downwelling, sky radiance
ε refers to emissivity of the surface pixel
τ refers to the atmospheric transmission

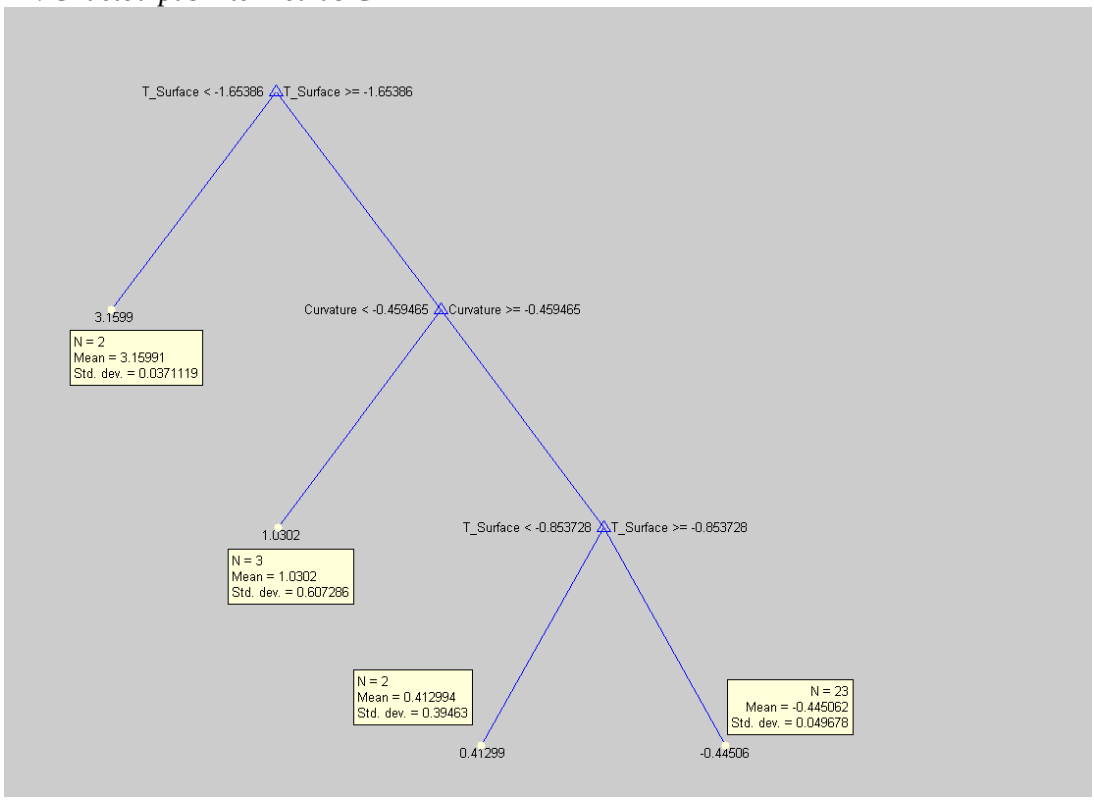
6. $T = k_2 / \ln((k_1/LT_\lambda) +1)$
where
T refers to the surface temperature (Kelvin)
k_1 and k_2 refer to Planck calibration constants for Landsat TM 5 band 6 (607.76 W/m^2 sr μm and 1260.56 K, respectively)
LT_λ refer to the surface blackbody radiance

Appendix 2: Regression trees for ten species of nocturnal desert rodents.

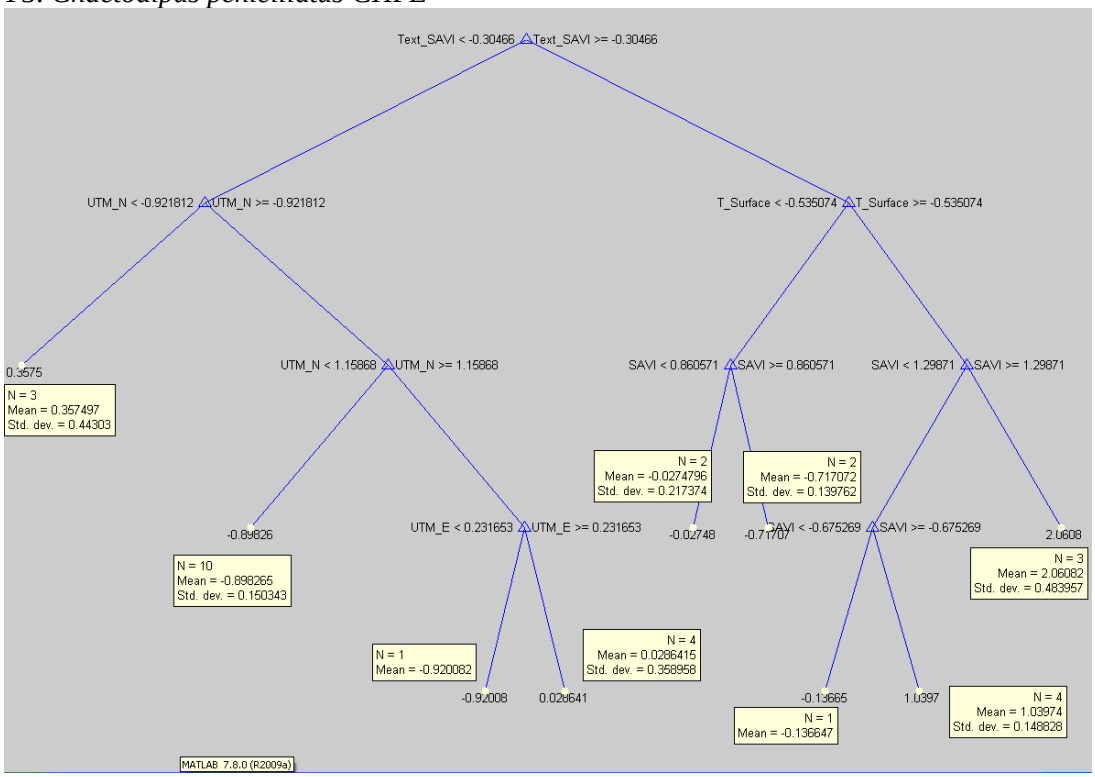
Y1: *Chaetodipus baileyi* CHBA



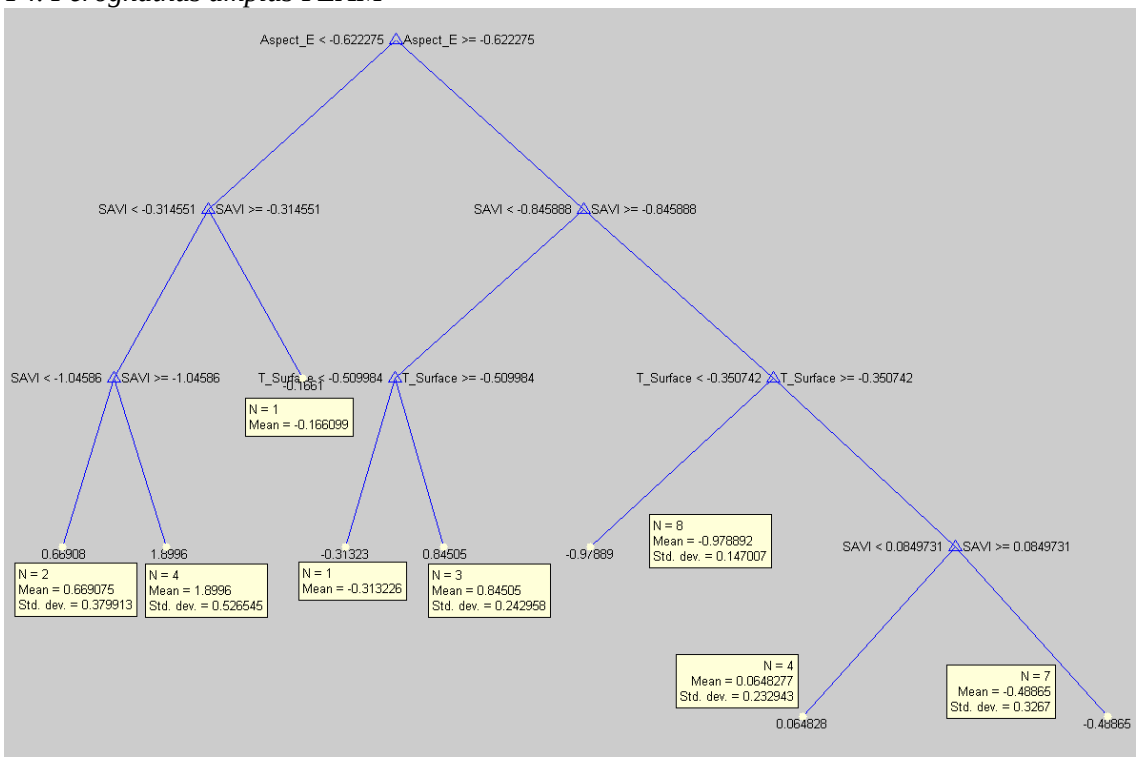
Y2: *Chaetodipus intermedius* CHIN



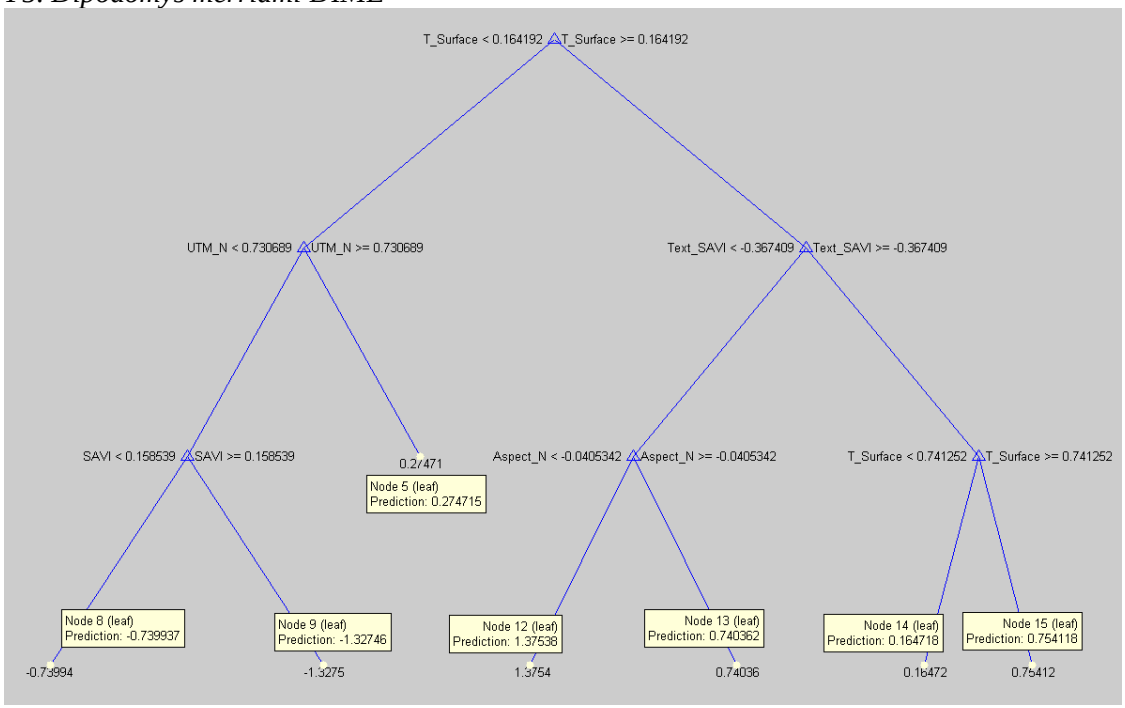
Y3: *Chaetodipus penicillatus* CHPE



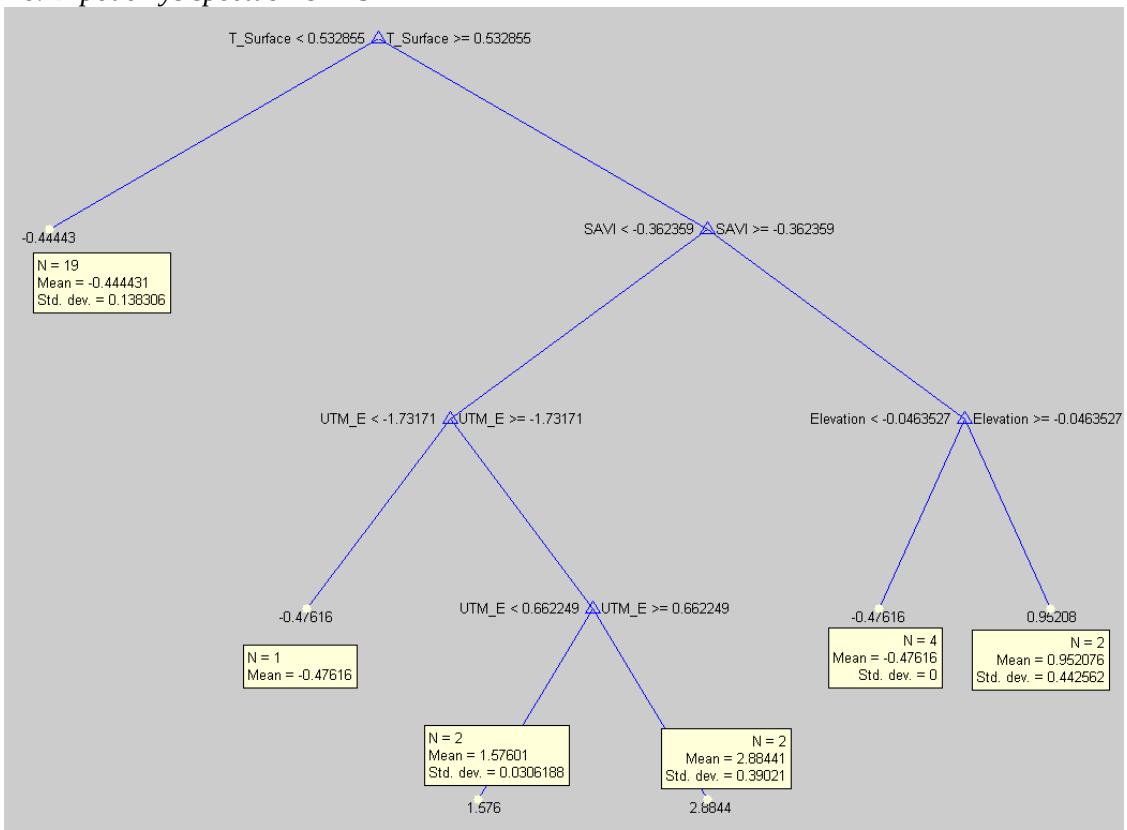
Y4: *Perognathus amplus* PEAM



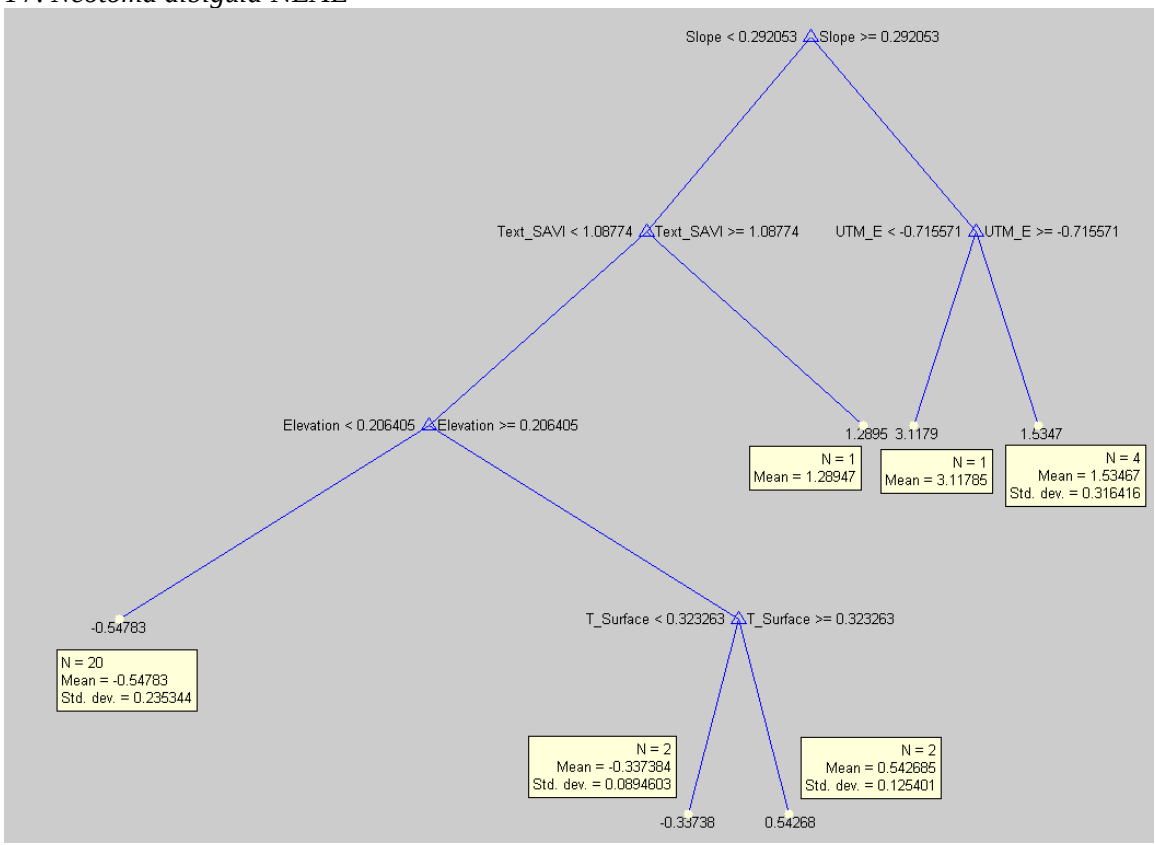
Y5: *Dipodomys merriami* DIME



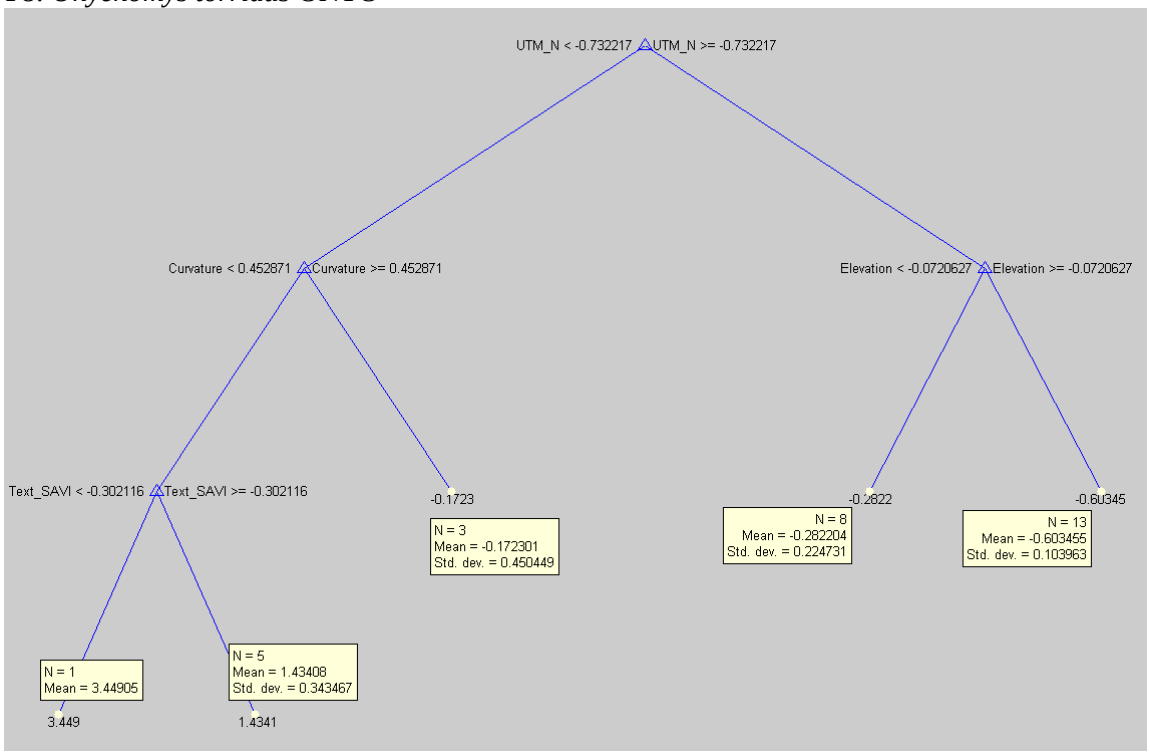
Y6: *Dipodomys spectabilis* DISP



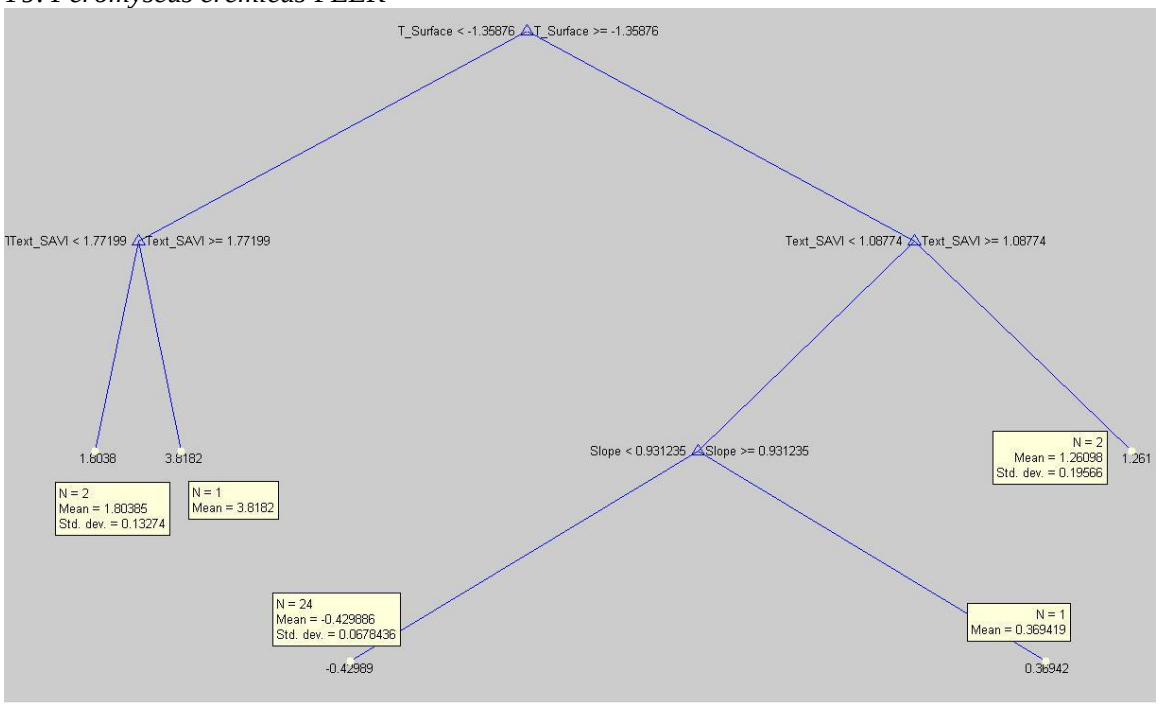
Y7: *Neotoma albigula* NEAL



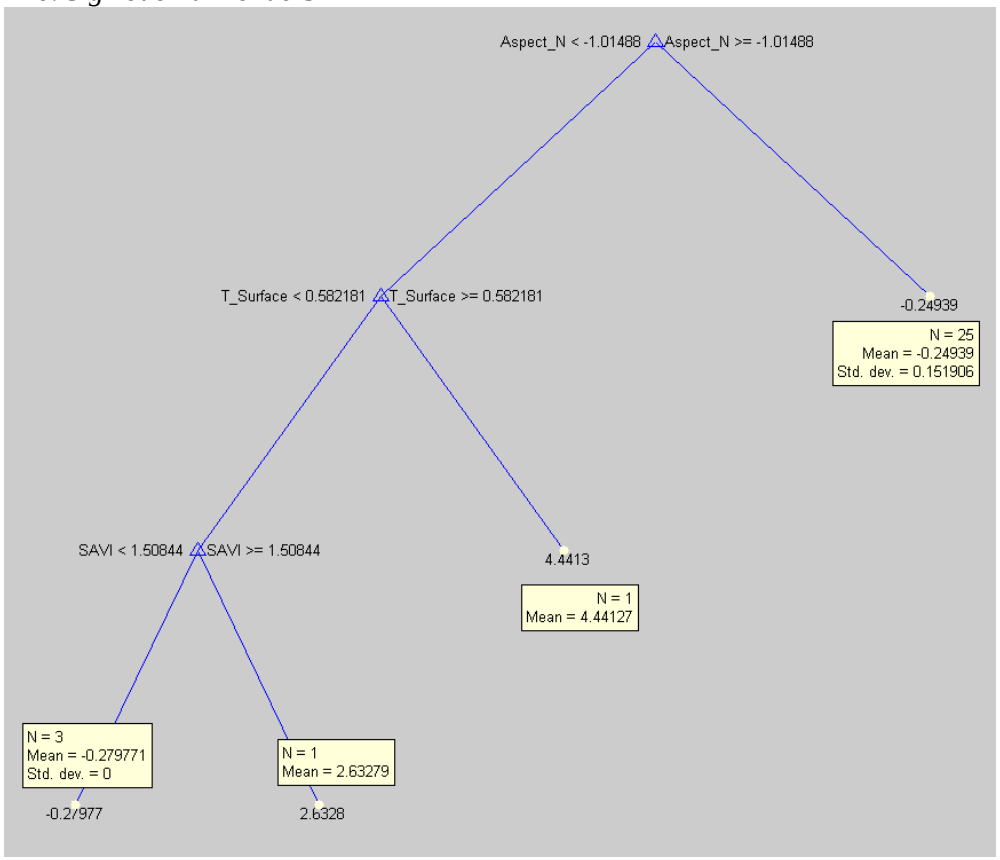
Y8: *Onychomys torridus* ONTO



Y9: *Peromyscus eremicus* PEER



Y10: *Sigmodon arizonae* SIAR



9. References

Akaike, H., 1973. Information Theory as an extension of the maximum likelihood principle. In: Petrov, B.N. and Csaki, F. (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest, Hungary, pp. 267-281.

Arizona-Sonora Desert Museum, 2000. A natural history of the Sonoran Desert, Phillips, S.J. and P.W. Comus, Eds., University of California Press, Berkeley CA, 650 pp.

Austin, M.P, Nicholls, A.O. and C.R. Margules, 1990. Measurement of the realized qualitative niche: environmental factors of several Eucalypt species in southern New South Wales. Ecological Monographs, 60: 161-177.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157:101-118.

Austin, M.P., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches, Ecological Modelling, 200, 1-19.

Beauvais, G.P and S.W. Buskirk, 1999. Modifying estimates of sampling effort to account for sprung traps, Wildlife Society Bulletin, 27(1): 39-43.

Barsi, J.A., Schott, J.R., Palluconi, F.D., and S.J. Hook, 2005. Validation of a Web-based atmospheric correction tool for single thermal band instruments. Proceedings of SPIE, Volume 5882, Earth Observing Systems X: 58820E-1, SPIE Bellingham WA, doi 10.1117/12.619990.

Boone, J.D., Otteson, E.W., McGwire, K.C., Villard, P., Rowe, J.E., and S.C. St. Jeor, 1998. Ecology and demographics of hantavirus infections in rodent populations in the Walker River Basin of Nevada and California, American Journal of Tropical Medicine and Hygiene 59(3): 445-451.

Brauner, N. and M. Shacham, 1998. Role of range and precision of the independent variable in regression of data. AIChE Journal, 44: 603-611.

Chamberlain, E., 1972. Soil survey, Organ Pipe Cactus National Monument, Pima County, Arizona. United States Department of Agriculture, Soil Conservation Service, 37p.

Chander, G., Markham, B. L., and D. L. Helder, 2009. Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors, Remote Sensing of Environment, 113: 893-903.

Chavez, P. S., Jr., 1996. Image-Based Atmospheric Corrections - Revisited and Improved, Photogrammetric Engineering & Remote Sensing, 62(9): 1025-1036.

Childs, J.E., Ksiazek, T.G., Pollesf, P.E., Krebs, J.W., and S. Zaki, 1994. Hanta viruses and their rodent reservoirs in the United States. Proceedings of the Sixteenth Vertebrate Pest Conference (1994), University of Nebraska-Lincoln.

De'ath, G. and K.E. Fabricius, 2000. Classification and regression trees, a powerful yet simple technique for ecological data analysis. Ecology, 81: 3178-3192.

Dormann, C.F., Purschke, O., Garcia-Marquez, J. Lautenbach, S. and B. Schroder, 2008. Components of uncertainty in species distribution analysis. Ecology 89: 3371-3386.

Dye, A.J., 1969. Germination potentials and accumulation of native plant seeds from southern New Mexico. New Mexico State University, M.S. Thesis.

Elith, J., Graham, C.H., Anderson, R.P., et al., 2006. Novel methods improve predictions of species distributions from occurrence data, Ecography, 29, 129-151.

Emmons, C.W., 1942. Isolation of Coccidioides from soil and rodents. Public Health Reports 57(4): 109-111.

ERDAS, Inc., Copyright © 1982 – 2008; Intergraph Corporation, http://geospatial.intergraph.com/products/ERDASIMAGINE/ERDASIMAGINE/Details.aspx.

Esri, 380 New York Street, Redlands, CA 92373-8100, USA, Copyright(C) 1995-2013 Esri
ArcGIS/ArcMap/ArcView Spatial Analyst

Franklin, J., 1995. Predictive vegetation mapping: geographical modeling of biospatial patterns in relation to environmental gradients, Progress in Physical Geography, 19: 474-499.

Franklin, J., and J.A. Miller, 2009. Mapping species distributions: spatial inference and prediction, Cambridge University Press, 2009, Cambridge, UK, 320 p.

Gage, K.L. and M.Y. Kosoy, 2005. Natural history of plague: perspectives from more than a century of research, Annual Review of Entomology 50:505-528,.

Goodchild, M.F., 1996. The spatial data infrastructure of environmental modeling. GIS and Environmental Modeling: Progress and Research Issues, Fort Collins, CO, GIS World Books, pp. 11-15.

Guisan, A. and N.E. Zimmermann, 2000. Predictive habitat distribution models in ecology, Ecological Modelling, 135: 147-186.

Guisan, A., Edwards, T.C., Jr., and T. Hastie, 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling, 157: 89-100.

Hastie, T., Tibshirani, R., and J. Freidman, 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction, 2nd Edition, New York, Springer-Verlag.

Henderson, P.A. and A.E. Magurran, 2010. Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. Proceedings of Biological Science, 277(1687): 1561-1570.

Hoffmeister, D.F., 1986. Mammals of Arizona, The University of Arizona Press, Tucson AZ, 600 p.

Holm, P., 2006. Nocturnal Rodents, Chapter 10, Organ Pipe Cactus National Monument Ecological Monitoring Report, 1997-2005. December 2006. National Park Service, U.S. Department of the Interior, Organ Pipe National Monument, Arizona http://www.nps.gov/orpi/naturescience/orpi-ecological-monitoring-report.htm

Huete, A. R. 1988. A Soil-adjusted Vegetation Index (SAVI), Remote Sensing of the Environment, 25:295-309.

Huete, A.R., Hua, G., Qi, J., Chehbouni, A. and W. J. Van Leeuwem, 1992. Normalization of Multidirectional Red and Near-infrared Reflectances with the SAVI, Remote Sensing of the Environment, 40:1-20.

Hulley, G.C. and S.J. Hook, 2009, The North American ASTER Land Surface Emissivity Database (NAALSED) Version 2.0, Remote Sensing of Environment, 113(9): 1967-1975.

Humes, K.S., Kustas, W.P., Moran, M.S., Nichols, W.D., and M.A. Weltz, 1994. Variability of Emissivity and Surface Temperature over a Sparsely Vegetated Surface, Water Resources Research, 30(5): 1299-1310.

Hunsaker, C.T., Nisbet, R.A., Lam, D.C.L, et al., 1993. Spatial models of ecological systems and processes: the role of GIS. Environmental Modeling with GIS, New York, NY: Oxford University Press, pp. 248-264.

Huston, M.A., 2002. Introductory essay: critical issues for improving predictions. In Scott, J.M., Heglund, P.J., Morrison, M.L., et al., (Eds.) Predicting Species Occurences: Issues of Accuracy and Scale. Covelo, CA: Island Press, pp.7-21.

Jiménez-Muñoz, J.C., Cristóbal J., Sobrino, J.A., Sòria, G., Ninyerola, M., Pons, X., 2009, Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. IEEE Transactions on Geoscience and Remote Sensing 47(1): 339-349.

Kadmon, R., Farber, O., and A. Danin, 2003. A systematic analysis of factors affecting the performance of climatic envelope models, Ecological Applications, 13: 853-867,

Kleiber, M., 1962. The Fire of Life, Wiley, New York NY.

Kochler, A.W. and I.S. Zonneveld, 1998. Vegetation Mapping. Boston: Kluwer Academic Publishers.

Kolivras, K.N. and A.C. Comrie, 2004. Climate and infectious disease in the southwestern United States, Progress in Physical Geography 28(3): 387-398,

Lawler, J.J., White, D., Neilson, R.P. and A.R. Blaustein, 2006. Predicting climate-induced range shifts: model differences and model reliability. Global Change Biology 12: 1568-1584.

Lazaroff, D.W., 1998. Arizona-Sonora Desert Museum Book of Answers, Arizona-Sonora Desert Museum Press, Tucson, AZ, 192 pp.

Legendre, P. and L. Legendre, 1998. Numerical Ecology, 2nd Ed. Elsevier Science BV, Amsterdam.

Loiselle, B.A., Jorgensen, P.M., Cosiglio, T., et al., 2008. Predicting species distribution from herbarium collections: does climate bias in sampling influence model outcomes?, Journal of Biogeography, 35:105-116.

Mateu, J., 1997. Methods of assessing and achieving normality applied to environmental data. Environmental Management, 21:767-777.

MATLAB Version 7.8.0.347 (R2009a), copyright 1984-2009, The MathWorks, Inc., www.mathworks.com.

Montandon, L.M. and E.E. Small, 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI, Remote Sensing of Environment, 112: 1835-1845.

Musick, H.B., 1975. Bareness of desert pavement in Yuma County, Arizona, Arizona Academy of Sciences Journal, 10(1): 24-28.

NAALSED, 2013. North American ASTER Land Surface Emissivity Database (NAALSED) https://lpdaac.usgs.gov/products/aster_products_table/naalsed.

NASA, 2012. "The Thematic Mapper, National Aeronautics and Space Administration," Washington D.C., http://landsat.gsfc.nasa.gov/about/tm.html.

NED, 2013. The National Elevation Dataset (NED), U.S. Department of Interior, U.S. Geological Survey, http://www.usgs.gov/pubprod/data.html#data.

Oindo, B.O. and A.K. Skidmore, 2002. Interannual variability of NDVI and species richness in Kenya, International Journal of Remote Sensing, 23: 285-298.

ORPI, 2006. Organ Pipe Cactus National Monument Ecological Monitoring Report, 1997-2005. December 2006. National Park Service, U.S. Department of the Interior, Organ Pipe National Monument, Arizona. http://www.nps.gov/orpi/naturescience/orpi-ecological-monitoring-report.htm

Petryszyn, Y., 1982. Population dynamics of nocturnal desert rodents: a nine year study, Dissertation, Department of Ecology and Evolutionary Biology, University of Arizona.

Petryszyn, Y., 1995. Small nocturnal mammals monitoring protocol for the Ecological Monitoring Program in Organ Pipe Cactus National Monument, Arizona. Organ Pipe Cactus National Monument Ecological Monitoring Program Protocol Manual, Special Report No. 11, National Biological Service Cooperative Park Studies Unit, The University of Arizona, Tucson, 5:1-23.

Petryszyn, Y., and E.L. Cockrum, 1990. Mammals of the Quitobaquito Management Area, Organ Pipe Cactus National Monument, Arizona, Technical Report No. 36, Cooperative National Park Resources Studies Unit, University of Arizona, Tucson, AZ, 32 p.

ReGAP, 2004. The Southwest Regional Gap Analysis Project, PROVISIONAL Digital Landcover Dataset for the Southwestern United States, Version 1.0, RS/GIS Laboratory, College of Natural Resources, Utah State University

Reichman, O.J., and M.V. Price, 1993. Ecological Aspects of Heteromyid Foraging. In Biology of the Heteromyidae, edited by H.H. Genoways and J.H. Brown. Special Publication No. 10 of the American Society of Mammalogists, 539–574.

Rosenzweig, M.L. and J Winakur, 1969. Ecology of desert rodent communities: habitats and environmental complexity, Ecology 50(4): 558-572.

Sakamoto, Y., Ishiguro, M., and G. Kitagawa, 1988. Akaike Information Criterion Statistics. KTK Scientific Publisher, Tokyo.

Sánchez, J.M., French, A.N., Mira, M., Hunsaker, D.J., Thorp, K.R., Valor, E., and V. Caselles, 2011, Thermal Infrared Emissivity Dependence on Soil Moisture in Field Conditions, IEEE Transactions on Geoscience and Remote Sensing, 49(11): 4652-4659.

Schott, J.R., Wiedman, A.B., and F. Wiedman, A Land Surface Temperature Product, http://landsat.usgs.gov/documents/LST%20Intro_short.pdf, Digital Image and Remote Sensing Laboratory, Rochester Institute of Technology, Rochester NY.

Scott, J.M., Heglund, P.J., Morrision, M.L. et al., Eds., 2002. Predicting Species Occurrences: Issues of Accuracy and Scale, Covelo CA, Island Press.

Scull, P, Franklin, J. and O.A. Chadwick, 2005. The application of classification tree analysis to soil type prediction in a desert landscape. Ecological Modelling, 181, 1-15.

Shreve, F. and I.L. Wiggins, 1964. Vegetation and flora of the Sonoran Desert. Stanford University Press, Palo Alto, 2 vols.

Sobrino, J.A., Jiménez-Muñoz, J.C., and L. Paolini, 2004. Land Surface Temperature Retrieval from LANDSAT TM 5, Remote Sensing of Environment 90: 434-440.

Sobrino, J.A., Jiménez-Muñoz, J.C., Sòria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., and P. Martínez, 2008. Land surface emissivity retrieval from different VNIR and TIR sensors, IEEE Transactions on Geoscience and Remote Sensing 48(2): 316-327.

SSURGO, 2012. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Arizona. Available online at http://soildatamart.nrcs.usda.gov .

Stamp, N.E. and R.D. Ohmart, Ecology, 59(4): 700-707.

Steenbergh, W.F. and P.L. Warren, 1977. Preliminary ecological investigation of natural community status at Organ Pipe Cactus National Monument. Cooperative National Park Research Studies Unit Technical Report No. 3, University of Arizona, Tucson, 152 p.

Saint-Germain, M., Buddle, C.M, Larrivee, M, Mercado, A. Motchula, T., Reichert, E., Sackett, T.E., Sylvain, Z., and A. Webb, 2007. Should biomass be considered more frequently as a currency in terrestrial arthropod community analysis? Journal of Applied Ecology 44: 330-339.

Stockwell, D.R.B., and A.T. Peterson, 2002. Effects of sample size on accuracy of species distribution models, Ecological Modelling, 148, 1-13.

Tabor J.A., 2009. Epidemiology of coccidioidomycosis in Tucson, Arizona, [dissertation]. Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, Arizona.

USGS, 2011. "Opening the Landsat Archive/Product Specifications," U. S. Geological Service, U.S. Department of the Interior, http://landsat.usgs.gov/products_data_at_no_charge.php.

Vayssieres, M.P, Plant, R.E. and B.H. Allen-Diaz, 2000. Classification trees, an alternative non-parametric approach for predicting species distributions. Journal of Vegetation Science, 11: 679-694.

Vincent, R.K., 1997. Fundamentals of Geological and Environmental Remote Sensing, Upper Saddle River, NJ, Prentice Hall.

Warren, P.L., Mortenson, B.K., Treadwell, B.D., Bowers, J.E. and K.L. Reichhardt, 1981. Technical Report No. 8, Vegetation of the Organ Pipe Cactus National Monument, Cooperative National Park Resources Studies Unit, University of Arizona/National Park Service

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California, Ecological Monographs, 30: 279-338.

Whittaker, R.H., 1967. Gradient analysis of vegetation. Biological Review 42: 207-264.

Wilson, J. and J. Gallant, 2000. Terrain Analysis: Principles and Applications, New York NY, John Wiley & Sons.

Wintle, B.A., Elith, J., and J.M. Potts, 2005, Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast region of NSW, Austral Ecology 30: 719-738.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., and A. Guisan, 2008. Effect of sample size on the performance of species distribution models, Diversity and Distributions, 14: 763-773.

Xiao, J., and A. Moody, 2005, A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA, Remote Sensing of Environment, 98: 237-250.

Yang, T.W. and C.H. Lowe, 1956. Correlation of major vegetation climaxes with soil characteristics in the Sonoran Desert, Science 123(3196): 542.

Zar, J.H. (1999). Biostatistical Analysis. 4th Edition. Prentice-Hall, New Jersey 931 pp.

Zeng, X., Dickinson, R.E., Walker, A., Shaikh, M., DeFries, R.S., and J. Qi, 2000, Derivation and Evaluation of Global 1-km Fractional Vegetation Cover Data for Land Modeling, Journal of Applied Meteorology, 39: 826-839.

APPENDIX B

FUGITIVE DUST EMISSION SOURCES ARISING FROM CONSTRUCTION: A REMOTE SENSING APPROACH

(TO BE SUBMITTED TO *GISCIENCE & REMOTE SENSING*)

ABSTRACT

Construction-related soil disturbance (e.g., road construction, trenching, landstripping, earthmoving and blasting), is a significant source of fugitive (airborne) dust in the atmosphere. Fugitive dust is a primary cause of decreased air quality and may carry airborne pathogens. We use Landsat Thematic Mapper (TM) remote sensing data spanning 1994 through 2009 over southern Arizona to identify source areas of construction-related activity likely to produce fugitive dust. We correlate temporal changes in the mid-infrared spectral response to dust sources from local construction. Image differencing of the TM band 5 (mid-infrared), with a change threshold of ± five standard deviations of the mean, suitably estimates the location and area affected by construction-related soil disturbance. Estimated dust-producing surface area ranges from 10.0 km^2 (1996-1997) to 28.3 km^2 (2004-2005), or 0.16% to 0.44% of the Pima County study area. Our methods aim to automate monitoring of fugitive dust sources by environmental and health agencies and to provide inputs to dust transport, air quality and climate models.

Keywords: soil disturbance, construction, fugitive dust, Valley Fever, remote sensing, Landsat, image differencing.

1. Introduction and Background

1.1 Introduction

Dust can be defined as "small solid particles, conventionally taken as those particles below 75 μm in diameter, which settle out under their own weight but which may remain suspended for some time" (ISO, 1994). Dust is termed *fugitive* when it is not discharged to the atmosphere in a confined flow stream, not collected by a capture system, or not emitted from a stack, chimney, or tailpipe (EPA, 1995). Common sources of fugitive dust include unpaved roads, bare ground, agricultural tilling, storage piles, and construction-related soil disturbance (Watson et al. 2002, Chow et al. 2003, Ho et al., 2003, Samara, 2005).

We aim in this paper to define space-time variations in fugitive dust, its health and environmental consequences and regulatory framework. We explore remote sensing technology to assess fugitive dust sources, introducing a change detection model to identify potential sources. Collateral dust inspection data and Geographical Information System (GIS) methods validate this model.

1.2 Health and environmental consequences of fugitive dust

Particles less than 10 μm are inhaled readily into the lungs where they may accumulate, react, be absorbed, or cleared. Scientific studies (EPA, 2003) have linked particle pollution with a series of significant health problems, including irritation of the airways, coughing, decreased lung function, aggravated asthma, chronic bronchitis, and premature death in people with heart or lung disease. Ecologists and environmental

scientists often overlook the importance of dust and wind-driven processes, yet these processes exert a fundamental influence on biogeochemical and ecological systems (Field et al, 2010).

As of August 30, 2011, EPA identified over 38 National Ambient Air Quality Standard (NAAQS) non-attainment areas in the U.S. for PM_{10} (particulate matter less than 10 μm), and these areas affect 39 counties with a population of over 25 million (EPA, 2012c). Large populations may therefore be exposed to unhealthy air, which in some cases is known to carry disease pathogens.

1.3 Coccidioidomycosis: A disease produced by the inhalation of airborne pathogens

Coccidioidomycosis (Valley Fever) is a systemic infection caused by inhalation of airborne spores from *Coccidioides spp*, which are fungi found in the soil in the southwestern U.S., including Arizona and California, and parts of Mexico (Galgiani, 1999). Coccidioidomycosis has increased epidemically in Arizona within the last decade (Park et al., 2005). The saprophytic phase of the species exists as slender filaments of cells that grow in the upper part of the soil (Kolivras et al., 2001).

Infection occurs typically following disturbance activities or natural events that disrupt the surface soil. This results in aerosolization of the fungal arthrospores and a subsequent spore source for inhalation by a host (Elconin et al., 1964, Maddy, 1957; Schneider et al., 1997; Swatek, 1970). There is a remarkable association of disease cases with previous exposures to dust storms, archeological digs, and occupational exposure to agricultural and construction dust (Emmons, 1942; Kirkland and Fierer, 1996;

Pappagianis and Einstien, 1978; Werner, 1974). The mechanisms of the association are not established. We look now at the origins of fugitive dust as a vehicle for pathogen dispersal.

1.4 Soil disturbance and construction-generated fugitive dust

A *disturbed surface area* is any portion of land that has been moved physically, uncovered, destabilized, or otherwise modified from its natural condition, thereby increasing the potential for fugitive dust emissions (EPA, 2012a). The dust-generation process during surface disturbance is caused by two basic physical phenomena: 1) pulverization and abrasion of surface materials by application of mechanical force through implements (wheels, blades, etc.), and 2) entrainment of dust particles by the action of turbulent air currents, such as wind erosion of an exposed surface by wind speeds over 19 kilometers per hour (km/hr, EPA, 1995).

Heavy construction, such as building and road construction, consists of a series of operations, for example, land clearing, frilling and blasting, ground excavation, earth moving, and structural construction (EPA, 1995). Each phase of the cycle exhibits its own duration and potential for dust generation. Furthermore, tracked out material on adjacent roadways can be suspended by traffic, and high wind events can lead to emissions from cleared land and material stockpiles.

1.5 Assessing the spatial and temporal distribution of fugitive dust sources

Fugitive dust sources arising from construction may be underestimated in typical assessments and modeling. Previous studies for fugitive dust sources focused mainly on

paved or unpaved road dust, soil dust and cement; source profiles for construction dust are limited (Kong et al., 2011).

Strategies for assessing and monitoring fugitive dust source areas include field surveys, collection of dust samples using various methods, census and statistical methods, and modeling of atmospheric circulation patterns (Stefanov et al., 2003). The multi-stage construction cycle also generates changes in spectral properties as the surface is modified (Jensen and Toll, 1982), which may be detected by remote sensing methods.

1.6 Remote sensing to detect change in surface spectral properties

Remote sensing, a source of multispectral data, can provide fundamental, new scientific information, including land cover change. The basic premise in using remote sensing data for change detection is that changes in land cover must result in changes in radiance values, and these changes must be large with respect to radiance changes caused by other factors such as atmospheric conditions, sun angle, and soil moisture (Ingram et al., 1981). Selecting the appropriate data and change detection technique according to the nature of the problem under investigation is critical in any change detection study (Jensen and Im, 2007).

Image differencing is the most widely applied algorithm for change detection (Coppin et al., 2004). Image differencing is the subtraction of spatially co-registered images collected at different times, producing a 'change' image (Jensen, 2005). Image subtraction produces positive and negative values in areas of radiance change, and values of zero in areas of no change. In an 8-bit remote sensing system (where pixel values range from 0 to 255), the potential range of difference values is -255 to +255. Identifying

appropriate spectral thresholds that discriminate “change” pixels from “no-change” pixels enables generation of a binary change image.

1.7 Image differencing to detect change at the urban fringe

Image differencing is known to provide superior results for identifying change due to natural or rural area conversion to urban land (Fung, 1990; Jensen and Toll, 1982, Ridd and Liu, 1998; Royal, 1980; Toll et al., 1980). This success is attributed partly to enhancement of contrast between vegetated and non-vegetated surfaces due to chlorophyll absorption, such as in the red portion of the spectrum.

None of the spectral band selection algorithms analyzed in these studies was found to be clearly superior; according to the authors, the choice depends on specific environmental conditions in play and application objectives (Ridd and Liu, 1998). The Landsat Thematic Mapper 5 (TM), which was not available for the earlier studies, provides data since 1984 in seven spectral regions, including the mid-infrared, at 30 meter spatial resolution (NASA, 2012b).

1.8 Mid-infrared spectral region for surface change detection

The TM middle infrared band 5 (1.55-1.75 μm) is known to discriminate moisture content of soil and vegetation. The spectral region is used often to monitor vegetation and soil, and it is used in drought and plant vigor studies (Howarth and Wickware, 1981; Maki, 2004; NASA, 2012a; USGS, 2010). The TM middle infrared (MIR) bands have been shown to account for much of the separability between wetland landcover types during image classification in one study (Jensen et al., 1993). TM band 5

(MIR) difference images were also used in this study to generate change/no-change masks of upland, non-tidal areas for use in classification preprocessing.

Generic spectral reflectance profiles suggest that soil, vegetation, and water are well-distinguished in the mid-infrared region, whereas spectral confusion can occur in other spectral regions. We compiled a list of soil and vegetation spectral indices which use the mid-infrared spectral region (Table 1). We focus on the Landsat mid-infrared for dust source detection studies in southern Arizona.

1.9 Southern Arizona study area

Our study area lies within the eastern portion of Pima County, Arizona, and consists of approximately 13 000 km² in area (Figure 1). Pima County is located in the southwestern region of the U.S and in the southern portion of Arizona. It shares approximately 200 km of border with the state of Sonora, Mexico. The population of Pima County is 980 263 (Census, 2012). The vast majority of the county population lies in and around the city of Tucson (518 956; 2006), which fills up much of the eastern part of the county. Tucson is Arizona's second largest city, and it is a major commercial and academic center. A ring of communities, unincorporated urban development, and undeveloped areas surround the city. The western half of the county is sparsely populated, and it is not included in the study area.

The Tucson region is a desert valley surrounded by five mountain ranges. The study area lies within the Sonoran Desert, which is a lush desert with legume trees and columnar cacti as dominant flora. Thousands of acres of the Sonoran Desert have been bladed for the construction of houses and commercial strip malls (AIA, 2007). Between

1980 and 1990, the city's area increased by over 50 percent through the annexation of unincorporated land, and currently, the total area encompassed within the city limits is 585 km^2. As the city has expanded, so too has development in the immediate 'fringe' areas.

2. DATA AND METHODS

2.1 Image preprocessing

Landsat Thematic Mapper 5 (TM) archive images are processed to the Level 1 Product Generation System with precision and terrain correction (LPGS, L1T), and with cubic convolution resampling (USGS, 2011). The eastern Pima County, AZ study area is located in Path 36 and Row 38 of the Landsat Worldwide Reference System. To minimize annual phenological effects, we selected Landsat 5 TM mid-infrared images captured in May or June for each year of the study period (1984-2009). Comparisons of random points within the area of interest showed good spatial registration to within one pixel between the images, and therefore additional fine-tuning beyond USGS processing was not required.

We incorporated updated radiometric calibration coefficients specific to the new USGS Landsat open-access archive (Chander et al, 2009) to achieve conversion of calibrated digital numbers to absolute units of at-sensor spectral radiance. We converted each image to atmospheric-corrected surface reflectance using the cosine approximation model (COST; Chavez, 1996). The COST model implements an improved dark-object atmospheric correction for Landsat TM multispectral data (bands 1-5 and 7).

2.2 Air quality dust inspection data

The Pima County Department of Environmental Quality (PDEQ) requires entities or individuals conducting land stripping or earthmoving in excess of one acre, trenching in excess of 300 feet, road construction in excess of 50 feet, and blasting to obtain an air quality activity permit through the PDEQ (Pima, 2012). In addition, the PDEQ operates an inspection program to monitor and regulate soil-disturbance activities likely to generate fugitive dust. The dust inspections are initiated by complaints from the public, referrals from other public agencies, and data obtained from the county's fugitive dust-generating activity permit program.

As of December of 2002, PDEQ inspectors document geographic coordinates for all PDEQ inspections using hand-held Global Positioning System (GPS) devices. More than 90% of permits received by the department are inspected, and each permit is usually inspected more than once to assess progress of the specific activity. The number of repeat inspections is determined by the respective activity's duration and scope (Pima, 2012).

We obtained data for PDEQ fugitive dust-generating activity inspections occurring between January 2003 and January 2006 by public information request. Spatial coordinates for each inspection point were subsetted into annual periods beginning in June and ending in May to match the satellite annual change detection image dates. The data were entered in a GIS, and used to develop and validate the remote sensing method.

2.3 Single-band image subtraction for change detection

We explored several techniques to ascertain change likely to be associated with soil disturbance. Based on simplicity, its association with soil properties and vegetation moisture (USGS, 2010; Table 1), and superior comparisons with the PDEQ dust inspection data set (see Discussion section 4.2), we selected single band differencing of the TM band 5 to perform our study.

Full-scene TM band 5 images were subtracted from the succeeding year's image to generate successive annual change images that cover the fifteen year span of the study (1994-2009). Each change image was subsetted to the approximate 640 km^2 study area. A mask was applied to remove areas in excess of 1000 meters elevation (generally the regional mountain ranges) in order to focus on the region's urban areas and their periphery. An additional mask derived the USGS Regional Southwest Gap Project (SWReGAP, 2004) was applied to remove agricultural and mining landcover in the northwest corner and south-center areas of the study area (10% of the remaining study area) to focus on construction-related activity.

Figure 2 displays a TM band 5 difference image of eastern Pima County for 2003-2004. Areas of greatest change are indicated by pixels with increased brightness (white) or increased darkness (black). Change resulting from anthropogenic disturbance is often identified in these images as having non-natural patterns, such as straight lines and rectangle, as highlighted by circles in the image.

2.4 Geographic information system analysis

We added the PDEQ dust inspection coordinates, subsetted to match the temporal periods of the satellite change images, into a GIS layer (ESRI ArcMap 9.2 and 9.3), and converted them to shape files. We applied a kernel density function to each set using consistent default input parameters, including a search radius (4670 meters) and an output grid cell size (0.220 square kilometers). The kernel density function, based on Silverman (1986), fits a surface over each point, with the surface value being highest at the location of the point and diminishes to zero at the search radius. The density at each output raster cell is calculated by adding the values of all the kernel surfaces where they overlay the raster cell center. Figure 3 displays the dust inspection data points and their corresponding kernel density plot for the 2004-2005 annual period.

We overlaid the satellite change images by their corresponding kernel density plot of dust inspection points. We are interested in the extent to which areas with high change detection signal (increased bright or dark pixels) are spatially consistent with areas showing high dust inspection density (reddish or darkish hue). The 2003-2004 comparison is displayed in Figure 4. Qualitative comparisons reveal good agreement between the PDEQ dust inspection and the remote sensing change detection data sets. We developed a thresholding strategy to limit false positives derived from these change detection protocols.

2.5 Fine-tuning the change thresholds

We initially established a change threshold of ± 4 standard deviations from the mean using visual inspection and co-spectral plots of the satellite change images with

their corresponding PDEQ dust inspection points and kernel density plots. Figure 5 displays the 2005-2004 change image presented as a binary image using the ± 4 standard deviation criterion. Pixels meeting the change threshold criterion are assigned black, and the corresponding PDEQ inspection kernel density plot overlain in grayscale.

To fine-tune selection of the band difference change threshold and to quantitatively assess model performance, we compared capture rates of thresholded change pixels by sets of buffers around the PDEQ dust inspection points to those of buffers around a set of equal number of random points. We used a 920 meter radius buffer size for all sets, which was derived from the average disturbance area reported in the PDEQ dust inspection data set (480 meters) with a factor of two in radius applied to ensure the inclusion of the full possible range of disturbance around any one dust inspection point. Buffer sets were generated for each of the four annual periods with PDEQ data (2002-2006), with an example shown in Figure 6.

We prepared a set of change images consisting of pixels thresholded from ±2.5 through 6.0 standard deviations from the mean, in increments of 0.5, for the four annual periods with dust inspection data. The capture rates or PDEQ buffers are compared with those of random points in Figure 7, with increasing threshold values.

3. RESULTS

3.1 Change threshold identification

The capture rate efficiency of thresholded change pixels by the PDEQ dust inspection buffers, compared to random buffers of the same size, increases as the standard deviation threshold increases, until approximately at ± five standard deviations

of the mean, where it plateaus (see Figure 7). We therefore selected ± five standard deviations of the mean as the optimum threshold.

Using this criterion, we classified the four change images of the annual periods with matching PDEQ data, and the average percentage of change pixels captured by the PDEQ dust inspection buffers was 50.13%, compared to 19.33% captured by the same number of random point buffers. Compared to random point buffers at 920 meters, each PDEQ dust inspection buffer captured statistically significantly more change pixels ($p < 0.0001$, Mann-Whitney rank sum *U*-test, difference in medians) for each of the comparisons. A summary of statistical comparisons is provided in Tables 2 and 3.

To test our choice of change detection method, we applied the threshold criterion, established for TM band 5 differencing, to products of several other change detection methods, including differencing of other Landsat bands, spectral index differencing and spectral transformation approaches. We subjected each output to the PDEQ and random buffer test. Figure 8 compares TM band 5 differencing with several other change detection methods for one annual period (2004-2003). At least at this criterion, TM band 5 differencing in the MIR is among the highest performers, each within a range of 38-42% of superior capture rate efficiency of PDEQ inspection buffers compared to random buffers.

Both qualitative and quantitative accuracy assessments show good cross-correlation between the Landsat TM change detection protocols and PDEQ dust inspection data. This is an important finding: although dust inspection permits go back only a few years, the Landsat TM time series extends back to the mid 1980s. We thus have an excellent historical record of surface disturbance for modeling purposes.

3.2 Change area determination

Change image pixels meeting the criterion of ±five standard deviations from the change image mean are classified as likely sources of construction-related soil disturbance based on their spatial and statistical association to fugitive dust inspection data. We use the output as a metric to estimate spatial and temporal aspects of the data set.

Using the dimensions of the TM band 5 pixel (900 square meters), we are able to estimate the surface area impacted by construction-related disturbance in the study area. The values for each annual period during the fifteen year span of the study are shown in Figure 9. Estimates ranged from10.0 km^2 (1996-1997) to 28.3 km^2 (2004-2005), or 0.16 - 0.44% of the masked study area, respectively. The model estimates that construction-related soil disturbance in eastern Pima County was fairly consistent before 2002, increased successively in the years 2002-2005 to a peak in 2005-2006, and decreased in recent years to pre-peak values.

We enumerated the thresholded change pixels to zip code spatial units with areal normalization for each annual period to assess sub-regional characteristics of the data set. Regional maps of three annual periods covering the span to the study period, with zip code areas experiencing increasing soil disturbance shown in increasing gray scale, are shown in Figure 10. We estimate locally high intensity of construction-related soil disturbance in the urban growth boundary zone of the peripheral north and northwest side of Tucson over the period of study, and increased local disturbance on the south side in the middle of the current decade.

4. DISCUSSION

4.1 Interfering activities

Surface changes generated from agriculture and mining activities such as crop rotation, field fallowing, and mining pond flooding, and range and forest fires, were outside the scope of the study. These activities generated significant change signals in some portions of the study area. Our approach was to mask out these areas; they affected only a small portion of study area and were easily isolated using ancillary landcover maps. Accurate ground-truth data, ancillary data, and additional data processing time and effort may be necessary to perform masking.

4.2 Comparisons to other work

We factor out seasonality using annual-date images, and we are thus provided with a yearly snapshot to capture the activity of a typical residential or commercial construction project cycle (6 months for one single family home or 11 months for a commercial site, EPA 1999). At this resolution, we are likely capturing the construction cycle and its dust-generating propensity as a whole, as opposed to characterizing individual stages. The 16-day repeat cycle for Landsat TM images provides numerous opportunities for scene selection.

Our use of image differencing to detect landcover change at the urban growth boundary is consistent with previous literature (Fung, 1990; Ridd and Liu, 1998; Royal, 1980; Toll et al., 1980). We report superior results using MIR differencing; many of the early studies emphasized use of the red spectral region. In our case, the pre-change

landcover of the Sonoran desert is anticipated to have a much stronger soil component than that of the other studies (i.e. grassland in Denver, CO or agriculture fields in Ontario, Canada and Salt Lake City, UT). We suggest that MIR is equally suited to detect the transformation of a pixel consisting of a significant soil component during disturbance and rearrangement, and subsequent construction.

An additional benefit of the MIR spectral region over visible spectral regions is its known resistance to the affects of atmospheric scattering. The COST model atmospheric adjustments we determined for the MIR images were consistently near zero; our results show that in the case of Landsat TM band 5, minimal preprocessing is necessary. Image differencing of a single spectral band requires minimal preprocessing compared to differencing of spectral indices and other spectral transformations such as principal component analysis.

We arrived at a rigorous image differencing change threshold (five standard deviations); previous work (Fung, 1990; Ridd and Liu, 1998; Quarmby and Cushnie, 1989) arrived at less stringent thresholds (0.7-1.2 standard deviations). We optimized our output to identify likely source areas of soil disturbance and fugitive dust generation; we can hypothesize that a more stringent criterion is needed to threshold change from a highly soil-dominated pixel characteristic of the Sonoran Desert to another form of soil (i.e., disturbed), than change from predominately agriculture or rangeland landcover. GIS-based modeling methods that test continuous thresholds and use genetic algorithm approaches to optimize moving threshold windows have been recently employed to automate and improve binary change detection performance (Im et al., 2008, 2009, 2011).

4.3 Assessment of accuracy

Our use of PDEQ fugitive dust inspection data with GIS techniques to guide development of the remote sensing method relies on its accuracy. This appears to be a good choice, but assumptions are made. Foremost is that the inspections point to the occurrence and location of dust generating activities. This is a straightforward assumption. In addition, we assume that the relationship between department inspections and dust generating activity is quantifiable. Dust inspections by PDEQ are quantifiable, as supported by discussions with PDEQ staff, in terms of repeat coverage based on duration and scope of the project at any particular site. Finally, we assume that the fugitive dust inspection program assesses all significant soil disturbance sources that generate dust. In this case, remote sensing methods may serve a role in assessing and validating administrative programs and policy.

Implicit in our method is that every pixel meeting threshold criteria equates to dust generating activity. We validate our approach using comparisons of the capture rates by buffers around know dust sources to those around random points. Based on our capture rates, we can say we are at least 50% accurate. In one sense, the model will likely overestimate the amount of area affected by soil disturbance, as not all change meeting criteria is due to construction, and not all construction phases are inherently dust-generating. The model may also underestimate this parameter, as change signals occurring on a surface area less than two times the pixel size of Landsat TM (30 meters), may not identified. Considering these limitations, the model may be best suited to compare *relative* change occurring year-to-year.

4.4 Potential applications of the output

Recent changes in policy and technology, including open access to the Landsat imagery archive, web browser tools and software that enable dynamic viewing of change imagery, and tools that perform auto-thresholding for change detection, have brought remote sensing change detection functionality to a broad range of users (Exelis, 2012; Green, 2011; Warner et al., 2009). Binary change images can provide information which is useful for planning and management purposes, especially when quick overview products are needed as preliminary output (Im et al., 2011). The 30m emission source metric reported here is in raster format, and it is easily converted to grid or GIS-based output at various scales and denominational units.

This output enables the assessment of the contribution of construction-related activities to total fugitive dust emissions, and it is a potential substitute or validation data for emission inventory data (EPA 2012b, 2012d). The characterization of the spatial density of emissions sources and their respective contributions (via emission factors) to fugitive dust provide a means for independent assessment of the efficiency of fugitive dust controls, abatement strategies and regulations. Similarly, we suggest that the output may also be useful for substitution and validation of methods employing source apportionment of PM monitoring network data. We, in addition, plan to use the output to characterize source areas of Coccidioides spp. spore propagation.

5. CONCLUSION

We employed remote sensing and GIS techniques to characterize construction-related soil disturbance likely to generate fugitive dust for an arid region of the U.S. Image differencing of the Landsat TM band 5 (mid-infrared), with a change threshold of ± five standard deviations of the mean, suitably estimates the location and area affected by construction-related soil disturbance in southern Arizona. We validated our method by comparison of capture rates of change pixels by buffers around known dust sources to those of an equal number of random points (50.13% vs. 19.33%, respectively, statistically significant difference). We estimate annual area affected by construction-related change, with values ranging from 10.0 km^2 (1996-1997) to 28.3 km^2 (2004-2005), or 0.16 to 0.44% of the masked study area, respectively. We aggregate this metric to zip code areas and identify regional hot spots and regional trends of the fifteen year study period.

The remote sensing approach described here is simple, requires minimal image preprocessing and processing, and makes use of readily accessible, free data. It locates areas of intense dust-generating activity for monitoring and mitigating fugitive dust. The spatially explicit output may serve as input or validation of models predicting the construction-related fugitive dust contribution to ambient air pollution, to perform atmospheric dispersion modeling, and to predict the dissemination of soil-borne pathogens such as the *Coccidioides* fungus.

6. ACKNOWLEDGEMENTS

Components of this research were funded with a grant from the Science to Achieve Results (STAR) program of the U.S. Environmental Protection Agency (#R8327540). Special thanks are extended to the Pima County Department of Environmental Quality for assistance with their dust inspection database and data collection methodology, and to the Arizona Regional Image Archive (ARIA) for assistance with COST image preprocessing protocol. Additional appreciation is extended to Dr. Daoqin Tong of the School of Geography and Development at the University of Arizona for assistance with statistical methods.

7. TABLES AND FIGURES

Table 1: Spectral indices emphasizing soil and vegetation moisture properties using Landsat TM. Spectral regions and the corresponding Landsat TM bands are shown in the *Formula* block.

Index	Formula	Reference
Bare Soil Index (BI)	[(SWIR + R) – (NIR + B)]/ [(SWIR + R) + (NIR +B)] (**B5**+B3)-(B4+B1)/(**B5**+B3)+(B4+B1)	Roy et al. (1997) Chen et al. (2004)
Normalized Difference Soil Index (NDSI)	[SWIR – NIR]/[SWIR + NIR] **B5**-B4/**B5**+B4	Rogers and Kearny (2004)
Normalized Difference Built-up Index (NDBI)	[SWIR – NIR]/[SWIR + NIR] **B5**-B4/**B5**+B4	Zha et al. (2003)
Built-Up Areas (BU)	NDBI - NDVI	He et al. (2010)
Normalized Difference Bareness Index (NDBaI)	[SWIR – TIR]/[SWIR + TIR] **B5**-B6/**B5**+B6	Zhao and Chen (2005)
Normalized Difference Water Index (NDWI)	[G – NIR]/[G + NIR] B2-B4/B2+B4	McFeeters (1996)
Normalized Difference Water Index (NDWI)	[R – SWIR]/[R + SWIR] B3-**B5**/B3+**B5**	Rogers and Kearny (2004)
Normalized Difference Water Index (NDWI)	[NIR – SWIR]/[NIR + SWIR] B4-**B5**/B4+**B5**	Gao (1996) Jackson et al. (2004)
Normalized Difference Moisture Index (NDMI)	[NIR – SWIR]/[NIR + SWIR] B4-**B5**/B4+**B5**	Wilson and Sader (2002)
Normalized Difference Infrared Index (NDII)	[NIR – SWIR]/[NIR + SWIR] B4-**B5**/B4+**B5**	Hardisky et al. (1983)

Table 2. Proportional comparison. Capture rates of satellite change pixels meeting threshold criteria (±five standard deviation) by 920m buffers around the Pima County Department of Environmental Quality (PDEQ) permit dust inspection points, as compared to buffers around the same number of random points, for June through May annual periods.

Year	Total DEQ or Random Points	Total Change Pixels	Pixels Inside 920m Buffers DEQ Points	Pixels Inside 920m Buffers Random Points
2003-2002	273[a]	14449	5148/14449 35.63%	1387/14449 9.60%
2004-2003	1043	19104	12945/19104 67.76%	5469/19104 28.63%
2005-2004	1013	31434	16504/31434 52.50%	9018/31434 28.69%
2006-2005	444[a]	21272	9492/21272 44.62%	2216/21272 10.42%
Average			50.13% (CV 0.27)	19.33% (CV 0.56)

[a] Partial year data sets: six months of inspection points only.

Table 3. Difference of means statistical tests. Example summary statistics for the 2005-2004 annual period comparing change pixel inclusion rates by PDEQ inspection point buffers (Group A) as compared to those of random point buffers (Group B).

Mann-Whitney *U* Test (non-parametric)	**Unpaired *t* Test**
$P < 0.0001$	$P < 0.0001$ (two-tailed; extremely statistically significant)
$U_A = 98581.5$	Difference of means = 69.10
$z = 31.48$	95% Confidence interval of difference: from 62.99 to 75.22
Mean Ranks 1422.7 vs. 604.3	Mean: 79.96 (DEQ buffers) vs. 10.86 (random point buffers)
$N_1 = N_2 = 1013$	SD: 86.55 vs. 48.19
	SEM: 2.72 vs. 1.51
	N: 1013

Figure 1. Study area map including the southwestern United States, northwestern Mexico, state of Arizona, Pima County and the city of Tucson.

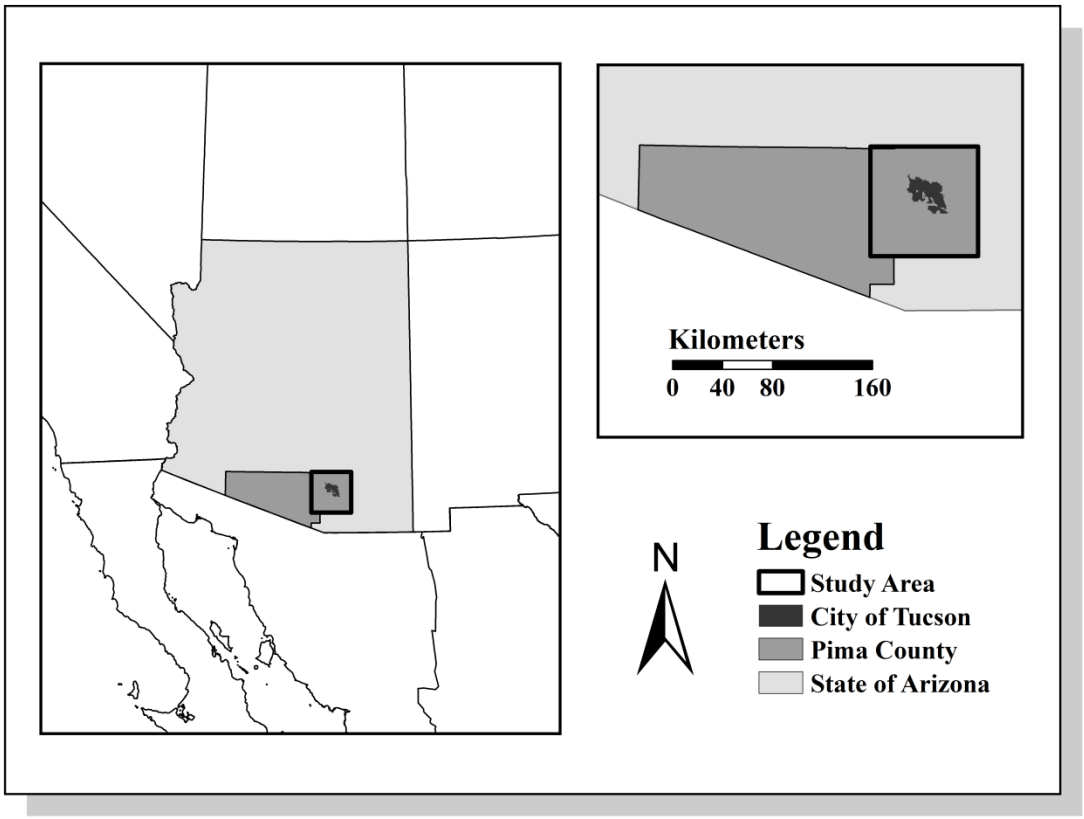


Figure 2. A change detection image of the Tucson, Arizona metropolitan region between May 2003 and June 2004 using band differencing in the mid-infrared spectral region (Landsat TM band 5). Change of interest to this study is noted as the bright or dark polygons throughout the image including the circled areas. This change is attributed primarily to human disturbance such as subdivision home construction. The large, bright area at the top (north) edge of the scene results from change induced by the Aspen fire occurring in the Santa Catalina Mountains during the summer of 2003. A mask was applied to remove areas in excess of 1000 meters elevation, thus the fire scar area was not included in the study.

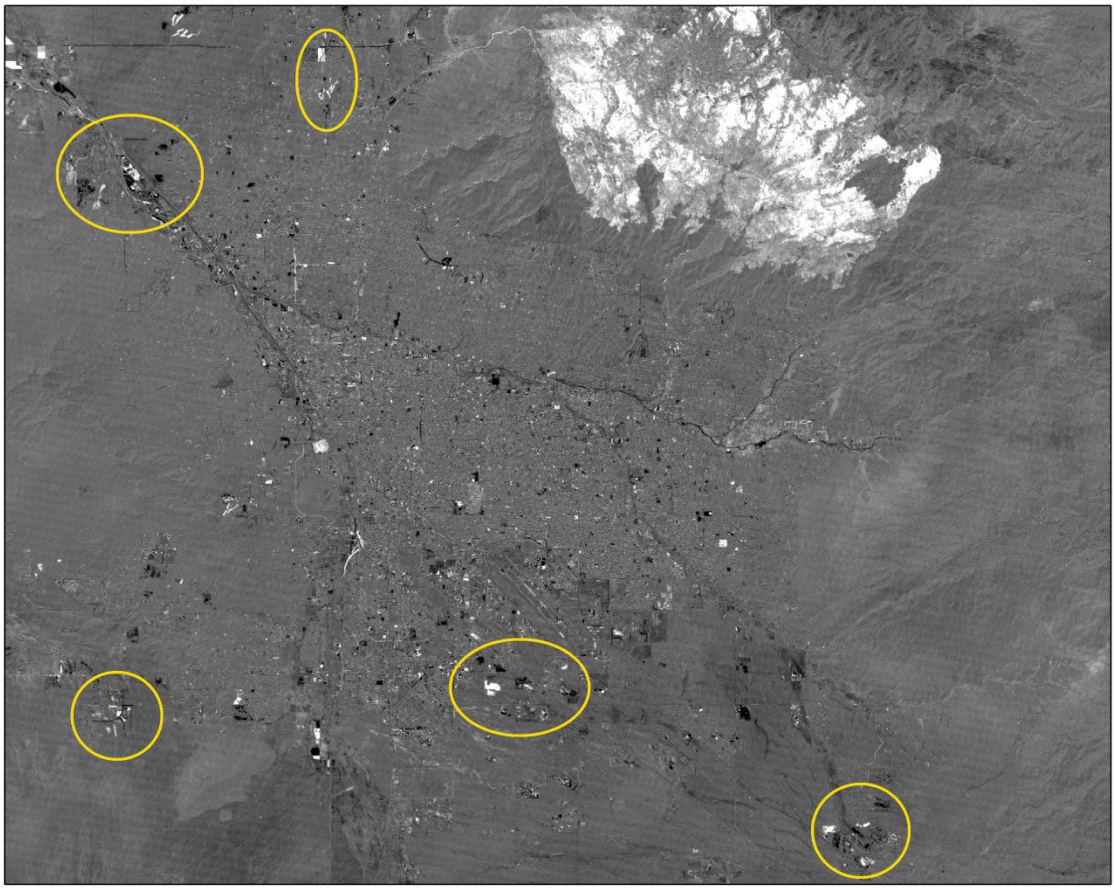


Figure 3. Spatial coordinates for fugitive dust inspections performed by the Pima County Department of Environmental Quality to assess construction-related activity between June 2004 and May 2005 are obtained and entered into a GIS (left). A kernel density function is applied to generate a density plot (right). A search radius of 4.67 kilometers and an output cell (grid) size of 0.220 square kilometers are used. Tucson city limits and the Pima county border are also shown.

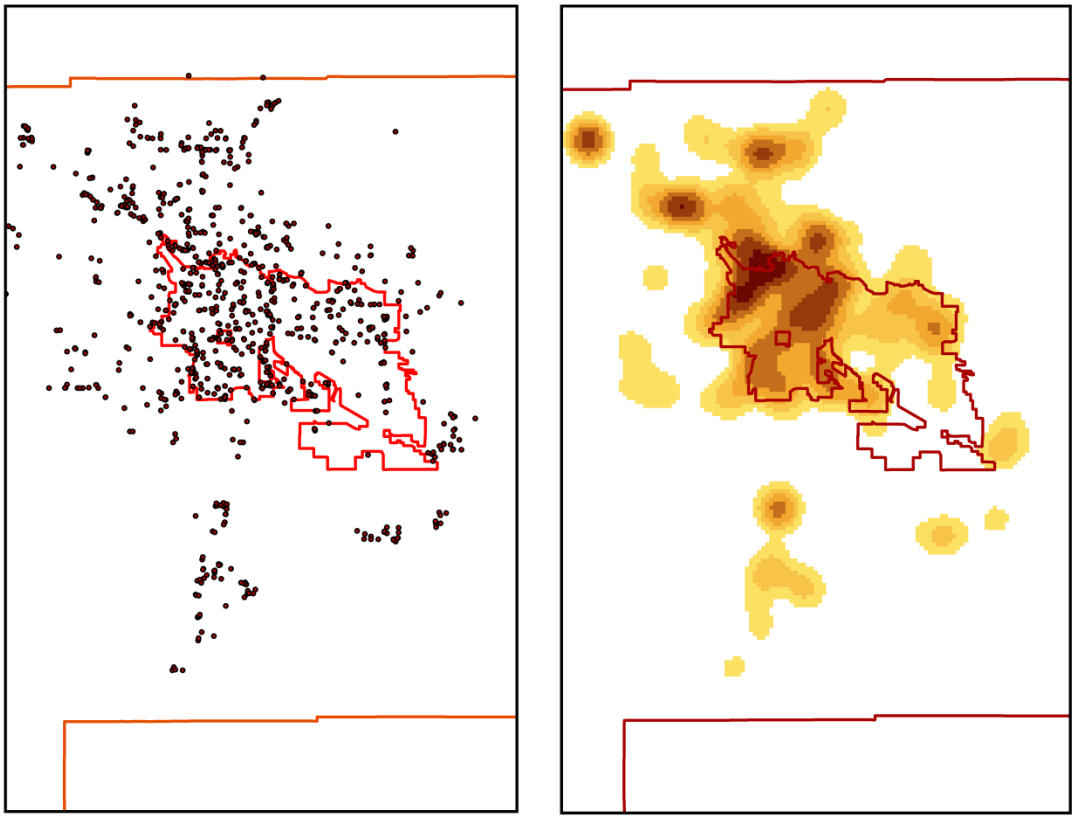


Figure 4. Remote sensing-derived change image with the Pima County DEQ dust inspection point kernel density overlay for the Tucson, Arizona metropolitan area between May 2003 and June 2004. Increased dust inspection density is indicated with increased reddish-brown or darkish hue, and change in the satellite change image is indicated by increased bright or dark pixels. We are interested in spatial alignment of dust inspection density and satellite change signal. The large bright feature in the north part of the image is due to change induced by the Aspen fire of 2003, and the bright and dark polygons in the lower left portion of the image are attributed to mining activities. Both of these areas are masked out before performing the study.

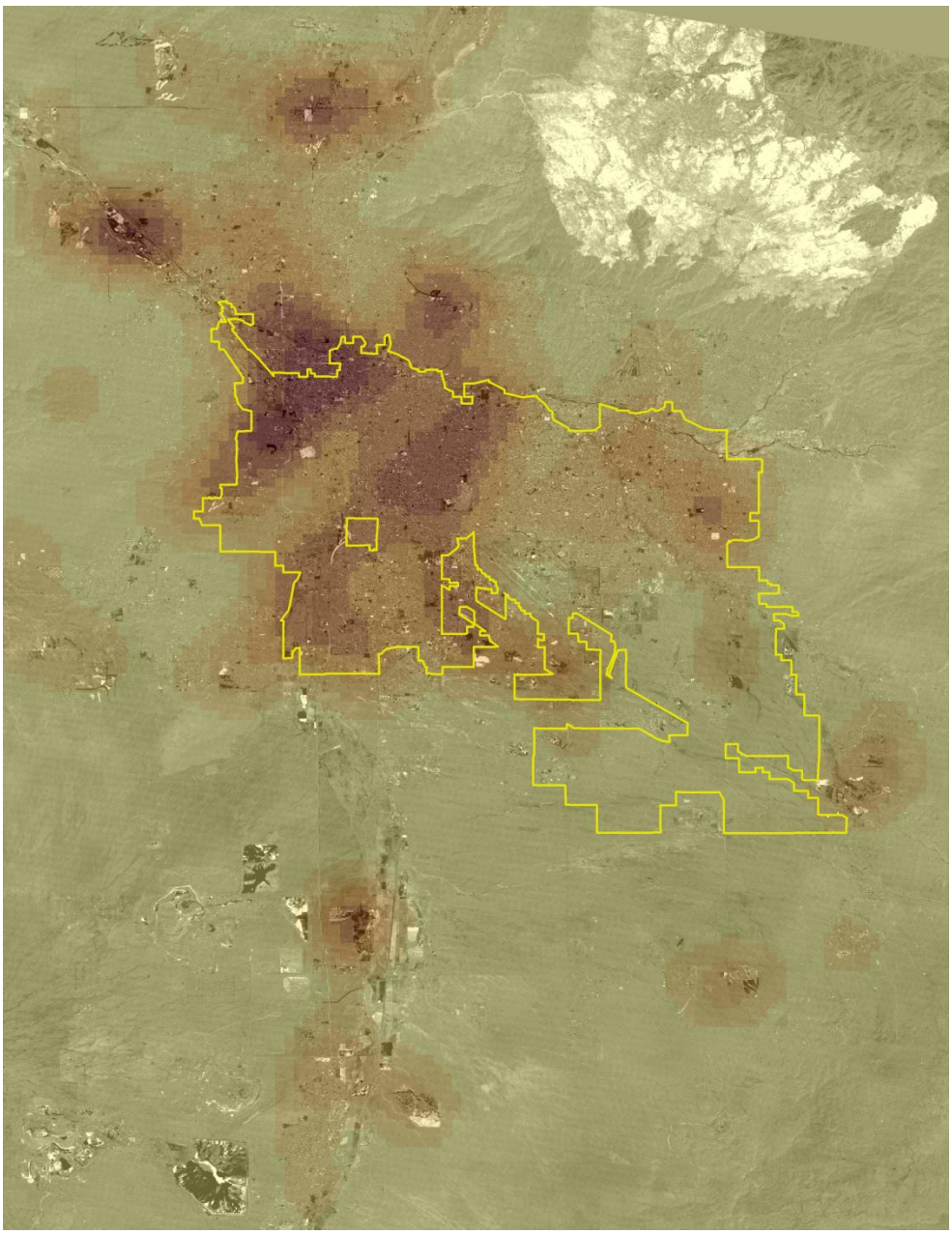


Figure 5. The remote sensing-derived change layer is converted to a binary image, with pixels exhibiting change values beyond ± four standard deviations displayed in black, and all other pixels displayed in white. The black pixels show greatest change signal between the two dates. The PDEQ dust inspection kernel density layer is shown in shades from white to dark gray with increasing density. Both data sets cover the period between June 2004 and May 2005.

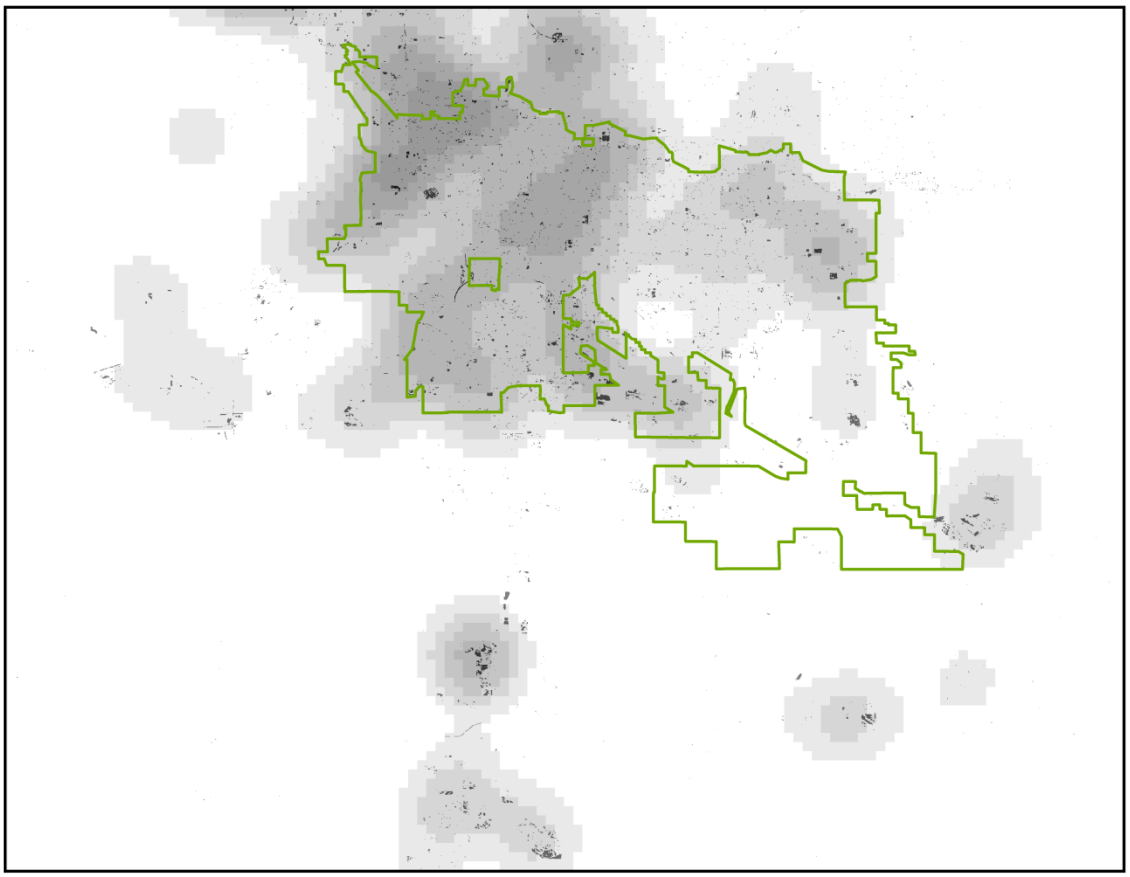


Figure 6. Demonstration of the buffer technique to assess model performance using the 2004-2003 data sets. Pixels meeting change criteria are shown as black polygons. 980 m buffers are shown in red (solid) lines around PDEQ dust inspection locations, and in blue (dotted) lines for random points. Capture rates of the change pixels are compared for the two buffer data sets to assess model performance.

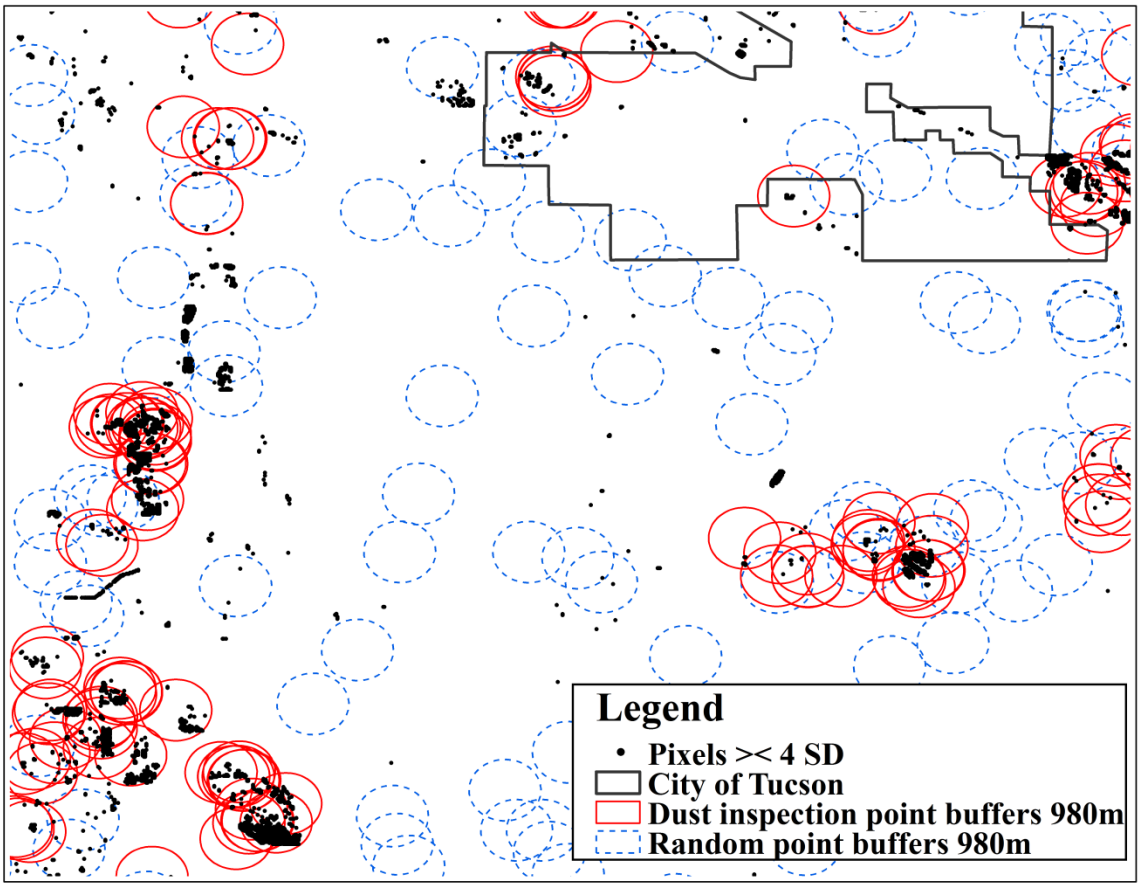


Figure 7. Proportion of Landsat TM band 5 image-differenced change pixels captured by buffers (920m) around county dust inspection points versus the proportion captured by an equal number of buffers around random points, compared at successively increasing standard deviation threshold. Values represent the average of four annual periods from 2003-2002 through 2006-2005.

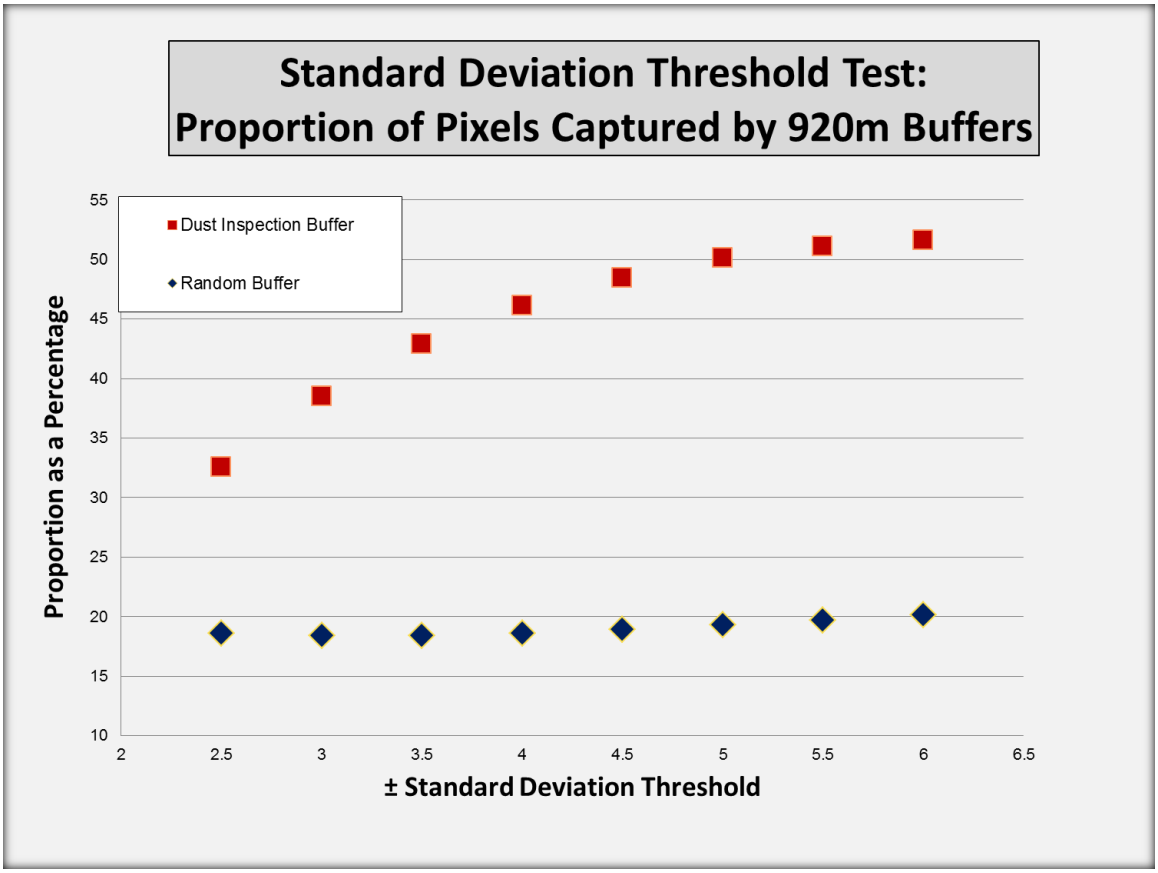


Figure 8. Comparison of change detection techniques. All methods are image differencing of Landsat TM bands or spectral indices, except for the spectral transformation approaches Principal Component Analysis (PCA) and Kauth-Thomas transformation layers. The difference between the percentage of change pixels meeting threshold criteria (± 5 standard deviation of the mean) that are captured by 920m buffers around a set of county dust inspection points to those captured by an equal number of random point buffer are shown. Data is compiled from the 2004 -2003 annual change image.

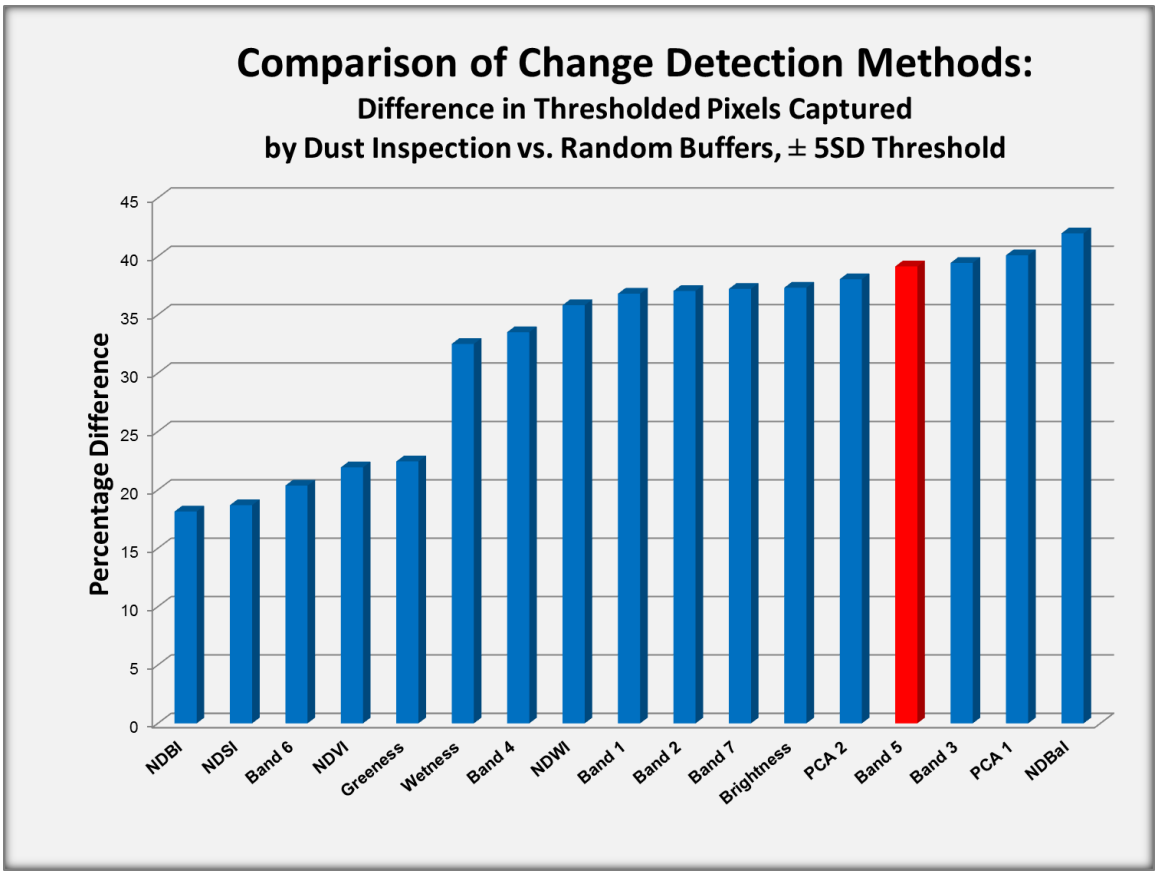


Figure 9. The aggregated area affected by construction-related soil disturbance liky to generarate fugitive dust in the study area for years between 1994 and 2009 is estimated. Disturbance areas are determined by summing pixels meeting threshold criteria: beyond ± five standard deviations of the mean in the Landsat TM band 5 difference image.

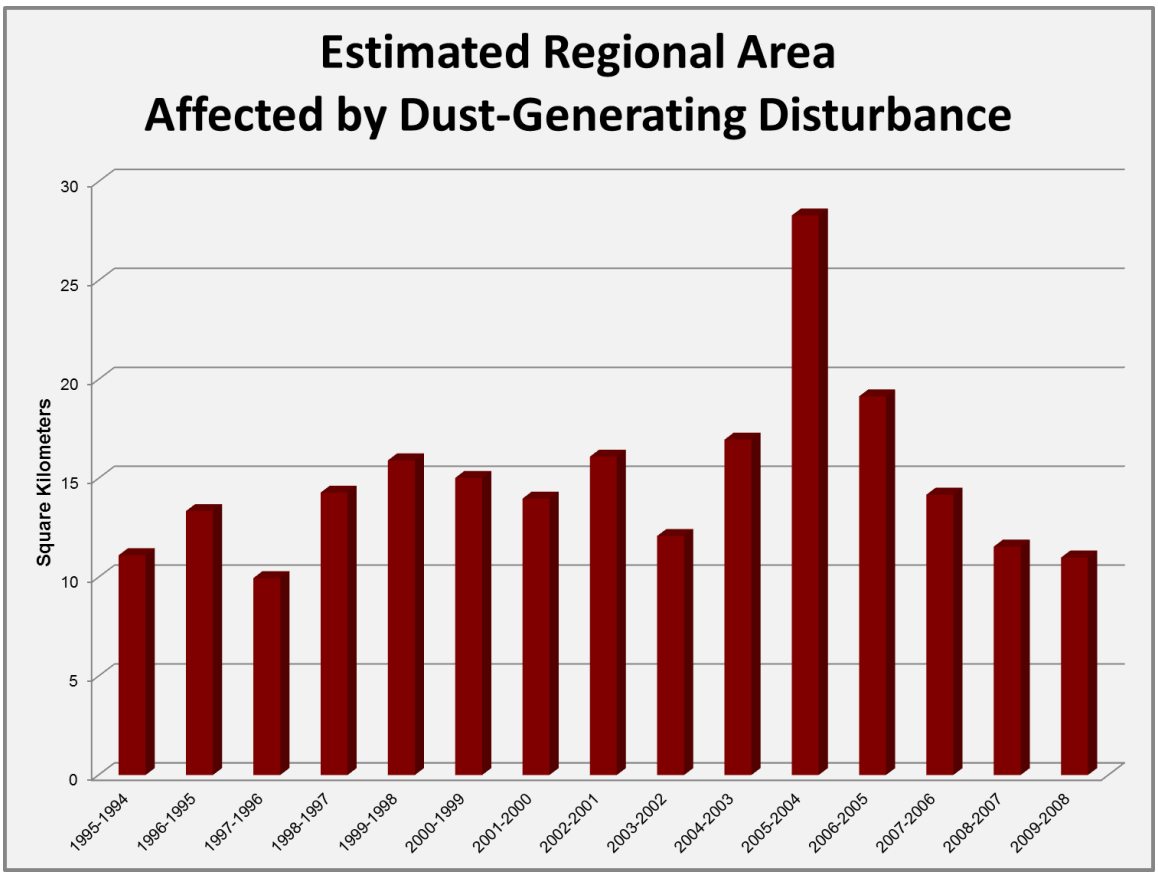
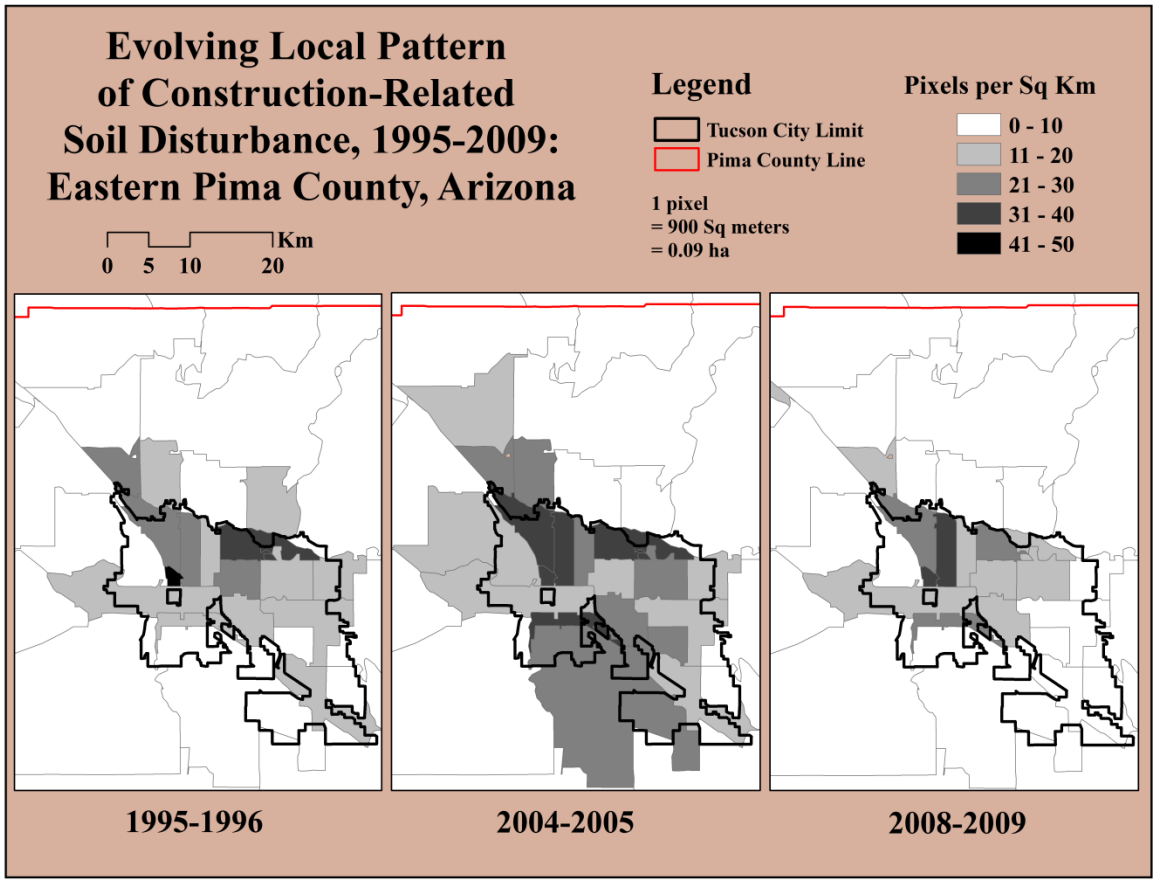


Figure 10. Remote sensing pixels meeting change criteria likely to indicate fugitive dust sources are aggregated to zip code areal units and normalized by area for three representative annual periods over the span of the study. Aggregation to smaller areal units enables locating hot spots of intense construction-related soil disturbance, and to monitor local trends in the region as a whole.



8. REFERENCES

AIA, 2007, "Tucson/Pima County, Arizona: One Million Reasons to Plan for Sustainable Growth," Southern Arizona Chapter of the American Institute of Architects, http://www.aia.org/aiaucmp/groups/aia/documents/pdf/aias078080.pdf.

Census, 2012, "State & County QuickFacts, Pima County, Arizona", U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/04/04019.html.

Chander, G., Markham, B. L., and D. L. Helder, 2009, "Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors," Remote Sensing of Environment, 113: 893-903.

Chavez, P. S., Jr., 1996, "Image-Based Atmospheric Corrections - Revisited and Improved," Photogrammetric Engineering & Remote Sensing, 62(9): 1025-1036.

Chen, W., Liu, L., Zhang, C., Wang, J., Wang, J., and Y. Pan, 2004, "Monitoring the Seasonal Bare Soil Areas in Beijing Using Multi-Temporal TM Images," Proceedings 2004 IEEE International, Geoscience and Remote Sensing Symposium, 5: 3379-3382.

Chow, J. C., Watson, J. G., Ashbaugh, L. L., and K. L. Magliano, 2003, "Similarities and Differences in PM10 Chemical Source Profiles for Geologic Dust from the San Joaquin Valley, California," Atmospheric Environment, 37(9-10): 1317-1340.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and E. Lambin, 2004, "Digital Change Detection Methods in Ecosystem Monitoring: A Review," International Journal of Remote Sensing, 25(9): 1565-1596.

Elconin, A. F., Egeberg, R. O., and M. C. Egeberg, 1964, "Significance of Soil Salinity on the Ecology of Coccidioides Immitis," Journal of Bacteriology, 87(3): 500-503.

Emmons, C. W., 1942, "Coccidioidomycosis," Mycologia, 34(4): 452-463.

EPA, 1995, "AP-42, Compilation of Air Pollutant Emission Factors," Chapter 13.2, Introduction to Fugitive Dust Sources, U.S. Environmental Protection Agency, http://www.epa.gov/ttn/chief/ap42/index.html#toc.

EPA, 1999, "Estimating particulate emissions from construction operations, Final Report," Prepared for the Emission Factor and Inventory Group, Office of Air Quality Planning and Standards, U.S. EPA, Research Triangle Park, NC, by the Midwest Research Institute, Kansas City, MO, EPA contract no. 68-D7-0068.

EPA, 2003, "Particle Pollution and Your Health," EPA-452/F-03-001, Office of Air and Radiation, U.S. Environmental Protection Agency, http://www.epa.gov/pm/pdfs/pm-color.pdf.

EPA, 2012a, "FIP Rule Definitions," U.S. Environmental Protection Agency, Pacific Southwest, Region 9, http://www.epa.gov/region9/air/phoenixpm/define.html.

EPA, 2012b, "The National Emissions Inventory: 2008 National Emissions Inventory Data," U.S. Environmental Protection Agency, Technology Transfer Network, Clearinghouse for Inventories & Emission Factors, http://www.epa.gov/ttn/chief/net/2008inventory.html.

EPA, 2012c, "The Green Book Nonattainment Areas for Criteria Pollutants," U.S. Environmental Protection Agency, http://www.epa.gov/airquality/greenbook/.

EPA, 2012d, "NEI Browser: County Emissions," U.S. Environmental Protection Agency, National Emissions Inventory, NEI Browser, http://neibrowser.epa.gov/eis-public-web/geo/county-emissions.html?stateJurisdictionId=4&inventoryYear=2008.

Exelis ENVI EX Image Differencing, 2012, ENVI EX Tutorial: Image Difference Change Detection, Boulder: Exelis Visual Information Solutions, 9 p.

Field, J. P, Belnap, J., Breshears, D. D., Neff, J. C., Okin, G. S., Whicker, J. J., Painter, T. H., Ravi, S., Reheis, M. C., and R. L. Reynolds, 2010, "The Ecology of Dust," Frontiers in Ecology and the Environment, 8(8): 423-430.

Fung, T., 1990, "An Assessment of TM Imagery for Land-Cover Change Detection," IEEE Transactions on Geoscience and Remote Sensing, 28(4): 681-684.

Galgiani, J. N., 1999, "Coccidioidomycosis: A Regional Disease of National Importance," Annals of Internal Medicine, 130(4): 293-300.

Gao, B., 1996, "NDWI- A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space," Remote Sensing of Environment, 58(3): 257-266.

Green, K., 2011, "Change Matters," Photogrammetric Engineering & Remote Sensing, 77(4): 305-309.

Hardisky, M. A., Smart, R. M., and V. Klemas, 1983, "Seasonal Spectral Characteristics and Aboveground Biomass of the Tidal Marsh Plant, Spartina alterniflora," Photogrammetric Engineering and Remote Sensing, 49(1): 85-92.

He, C., Shi, P., Xie, D., and Y. Zhao, 2010, "Improving the Normalized Difference Built-Up Index to Map Urban Built-Up Areas Using a Semiautomatic Segmentation Approach," Remote Sensing Letters, 1(4): 213-221.

Ho, K. F., Lee, S. C, Chow, J. C., and J. G. Watson, 2003, "Characterization of PM10 and PM2.5 Source Profiles for Fugitive Dust in Hong Kong," Atmospheric Environment, 37(8): 1023-1032.

Howarth, P. J. and G. M. Wickware, 1981, "Procedures for Change Detection Using Landsat Digital Data," International Journal of Remotes Sensing, 2(3): 277-291.

Im, J., Rhee, J., Jensen, J. R. and M. E. Hodgson, 2008, "Optimizing the Binary Discriminant Function in Change Detection Applications," Remote Sensing of Environment, 112(2008): 2761-2776.

Im, J., Rhee, J. and J. R. Jensen, 2009, "Enhancing Binary Change Detection Performance Using a Moving Threshold Window (MTW) Approach," Photogrammetric Engineering & Remote Sensing, 75(8): 951-962.

Im, J., Lu, Z. and J. R. Jensen, 2011, "A Genetic Algorithm Approach to Moving Threshold Optimization for Binary Change Detection," Photogrammetric Engineering & Remote Sensing, 77(2): 167-180.

Ingram, K., Knapp, E., and J. W. Robinson, 1981, "Change Detection Technique Development for Improved Urbanized Area Delineation," Technical Memorandum, CSC/TM-81/6087, Computer Sciences Corporation, Silver Spring MD, p. 75.

ISO, 1994, "Air Quality – General Aspects, Vocabulary," in: ISO Standard 4225, International Organization for Standardization (ISO), ISO TC 146/SC 4, Geneva.

Jackson, T. J., Chen, D., Cosh, M., Fuqin, L., Anderson, M., Walthall, C., Doriaswamy, P., and E. R. Hunt, 2004, "Vegetation Water Content Mapping Using Landsat Data Derived Normalized Difference Water Index for Corn and Soybeans," Remote Sensing of Environment, 92(4): 475-482.

Jensen, J. R. and D. L. Toll, 1982, "Detecting Residential Land-Use Development at the Urban Fringe," Photogrammetric Engineering and Remote Sensing, 48(4): 629-643.

Jensen, J. R., Cowen, D., Althausen, J., Narumalani, S. and O. Weatherbee, 1993, "An Evaluation of CoastWatch Change Detection Protocols in South Carolina," Photogrammetric Engineering & Remote Sensing, 59(6): 1039-1046.

Jensen, J. R., 2005, Introductory Digital Image Processing: A Remote Sensing Perspective, (3rd Ed.), Upper Saddle River: Prentice-Hall, 526 p.

Jensen, J. R. and J. Im., 2007, "Remote Sensing Change Detection in Urban Environments," in R.R. Jensen, J. D. Gatrell and D. D. McLean (Eds.), Geo-Spatial Technologies in Urban Environments Policy, Practice, and Pixels, (2nd Ed.), Berlin: Springer-Verlag, 7-32.

Kirkland, T.N. and J. Fierer, 1996, "Coccidioidomycosis: A Reemerging Infectious Disease," Emerging Infectious Diseases, 3(2): 192-199.

Kolivras, K. N., Johnson P. S., Comrie, A. C. and S. R. Yool, 2001, "Environmental Variability and Coccidioidomycosis (Valley Fever)," Aerobiologia, 17: 31-42.

Kong, S., Ji, Y., Lu, B., Chen, L., Han, B., Li, Z., and Z. Bai, 2011, "Characterization of PM10 Source Profiles for Fugitive Dust in Fushun-A City Famous for Coal," Atmospheric Environment, 45: 5351-5365.

Maddy, K., 1957, "Ecological Factors of the Geographic Distribution of Coccidioides immitis," Journal of the American Veterinary Medical Association, 130: 475-476.

Maki, M., Ishiahra, M., and M. Tamura, 2004, "Estimation of Leaf Water Status to Monitor the Risk of Forest Fires by Using Remotely Sensed Data," Remote Sensing of Environment, 90: 441-450.

McFeeters, S. K, 1996, "The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features," International Journal of Remote Sensing, 17(7): 1425-1432.

NASA 2012a, "Remote Sensing Overview Presentation," U. S. Geological Service, U.S. Department of the Interior, http://l7downloads.gsfc.nasa.gov/assets/other/RSI_notes.pdf.

NASA, 2012b, "The Thematic Mapper, National Aeronautics and Space Administration," Washington D.C., http://landsat.gsfc.nasa.gov/about/tm.html.

Pappagianis, D, and H. Einstein, 1978, "Tempest from Tehachapi Takes Toll or Coccidioides Conveyed Aloft and Afar," Western Journal of Medicine, 127: 527-530.

Park. J. B., Siegel, K., Vaz, V., Komatsu, K., McRill, C., Phelan, M., Colman, T., Comrie, A. C., Warnock, D. W., Galgiani, J. N., and R. A. Hajjeh, 2005, "An Epidemic of Coccidioidomycosis in Arizona Associated with Climatic Changes, 1998-2001," The Journal of Infectious Diseases, 191: 1981-1987.

Pima, 2012, "Fugitive Dust FAQs," Pima County, Arizona, http://www.deq.pima.gov/air/FugitiveDustFAQs.htm.

Quarmby, N. A. and J. L. Cushnie, 1989, "Monitoring Urban Land Cover Changes at the Urban Fringe from SPOT HRV Imagery in South-East England," International Journal of Remote Sensing, 10(6): 953-963.

Ridd, M. K. and J. Liu, 1998, "A Comparison of Four Algorithms for Change Detection in an Urban Environment," Remote Sensing of Environment, 63: 95-100.

Rogers, A. S. and M. S. Kearney, 2004, "Reducing Signature Variability in Unmixing Coastal Marsh Thematic Mapper Scenes Using Spectral Indices," International Journal of Remote Sensing, 25(12): 2317-2335.

Roy, P. S., Miyatake, S. and A. Rikimaru, 1997, "Biophysical spectral response modeling approach for forest density stratification," ACRS1997 Proceedings, Asian Conference on Remote Sensing of the Environment, FR97-7, Kuala Lumpur.

Royal, J.A., 1980, "Change Detection Method Development Census Urban Area Application Pilot Test," Final Report, Contract NAS5-25707, General Electric Company, Beltsville MD, 74 p.

Samara, C., 2005, "Chemical Mass Balance Source Apportionment of TSP in a Lignite-Burning Area of Western Macedonia, Greece," Atmospheric Environment, 39(34): 6430-6443.

Schnieder, E. et al., 1997, "A Coccidioidomycosis Outbreak Following the Northridge, Calif, Earthquake," The Journal of the American Medical Association, 277(11): 904-908.

Silverman, B. W., 1986, Density Estimation for Statistics and Data Analysis, New York: Chapman and Hall, 176 p.

Stefanov, W. L., Ramsey, M. S., and P. R. Christensen, 2003, "Identification of Fugitive Dust Generation, Transport, and Deposition Areas Using Remote Sensing," Environmental & Engineering Geoscience, IX(2): 151-165.

Swatek, F. E., 1970, "Ecology of Coccidioides Immitis," Mycopathologia et Mycologia Applicata, 40: 3-12.

SWReGAP, 2004, "Provisional Digital Land Cover Map for the Southwestern United States," Version 1.0, USGS National Gap Analysis Program, RS/GIS Laboratory, College of Natural Resources, Utah State University, http://earth.gis.usu.edu/swgap/landcover.html.

Toll, D. L., Royal, J. A. and J. B. Davis, 1980, "Urban Area Update Procedures Using Landsat Data," Proceedings of the American Society of Photogrammetry, ASP Fall Technical Meeting, RS-E-1-17, 12 p.

USGS, 2010, "Frequently Asked Questions about the Landsat Missions," U. S. Geological Service, U.S. Department of the Interior, http://landsat.usgs.gov/best_spectral_bands_to_use.php.

USGS, 2011, "Opening the Landsat Archive/Product Specifications," U. S. Geological Service, U.S. Department of the Interior, http://landsat.usgs.gov/products_data_at_no_charge.php.

Warner, T. A., Almutairi, A. and J. Y. Lee, 2009, "Remote Sensing of Land Cover Change," in Warner, T. A., Nelllis, M. D. and G. M. Foody (Eds.), The Sage Handbook of Remote Sensing, Los Angeles, Sage, Inc., 459-472.

Watson, J. G., Zhu, T., Chow, J. C., Engelbrecht, J., Fujita, E. M. and W. E. Wilson, 2002, "Receptor Modeling Application Framework for Particle Source Apportionment," Chemosphere, 49(9): 1093-1136.

Werner, S. B., 1974, "Coccidioidomycosis Among Archeology Students: Recommendations for Prevention," American Antiquity, 39(2): 367-370.

Wilson, E. H. and S. A. Sader, 2002, "Detection of Forest Harvest Type Using Multiple Dates of Landsat TM Imagery," Remote Sensing of Environment, 80(3): 385-396.

Zha, Y., Ga, J. and S. Ni, 2003, "Use of Normalized Difference Built-Up Index in Automatically Mapping Urban Areas from TM Imagery," International Journal of Remote Sensing, 24(3): 583-594.

Zhao, H. and X. Chen, 2005, "Use of Normalized Difference Bareness Index in Quickly Mapping Bare Areas from TM/ETM+," IEEE International, Proceedings IGARSS '05, International Geoscience and Remote Sensing Symposium, 0-7803-9050-4, 10.1109/IGARSS.2005.1526319: 1666-1668.

APPENDIX C

ESTIMATING ENVIRONMENTAL SOURCES OF VALLEY FEVER PROPAGATION IN SOUTHERN ARIZONA WITH REMOTE SENSING

(TO BE SUBMITTED TO *EMERGING INFECTIOUS DISEASES*)

ABSTRACT

Coccidioidomycosis (Valley Fever) is an environmentally-mediated respiratory disease caused by the inhalation of airborne spores from the fungi *Coccidioides spp*. The fungi reside in arid and semi-arid soils of the Americas. The disease has increased epidemically in Arizona and other areas within the last decade. Despite this increase, the ecology of the fungus remains obscure, and environmental antecedents of the disease are largely unstudied. We investigate two sources of soil disturbance, anthropogenic construction and nocturnal desert rodents, which are hypothesized to affect soil ecology and initiate *Coccidioides spp.* spore dissemination. We estimate construction-related soil disturbance with annual differencing of Landsat Thematic Mapper mid-infrared images. Source areas of soil disturbance are identified, and annual affected areas are estimated for eastern Pima County, Arizona and for zip code areas spanning 1994 through 2006. We build rodent abundance distribution maps for the study area using regression models of biophysical variables derived from remotely-sensed data with comparisons to rodent trapping data from the Organ Pipe Cactus National Monument. The two spatially explicit soil disturbance sources are compared with coccidioidomycosis incidence data using rank order correlation and regression methods. Construction-related soil disturbance correlates with annual county-wide incidence (R^2 = 0.49, p-value 0.012), and with incidence of zip codes at the periphery of the city of Tucson for the total study period (R^2 = 0.48, p-value 0.001). The average abundance values for the desert pocket mouse (*Chaetodipus penicillatus*), derived from a soil-adjusted vegetation index, aspect (northing) and thermal radiance, correlate with total study period incidence aggregated to zip code (R^2 = 0.25, p-value 0.02).

Key Words: coccidioidomycosis, Valley Fever, remote sensing, Landsat, rodent, species distribution modeling, Sonoran Desert, construction, soil disturbance, medical geography, Arizona.

1. Introduction

1.1 Coccidioidomycosis

Coccidioidomycosis (Valley Fever) is a systemic infection characterized by fever, respiratory infection and reddish bumps on the skin. It is caused by inhalation of airborne spores from *Coccidioides immitis* and *Coccidioides posadasii* fungi, which are endemic in the southwestern United States, and in parts of Mexico, Central and South America. Coccidioidomycosis is produced in Arizona by inhalation of spores from *C. posadasii* (Fisher et al., 2002).

The *Coccidioides spp.* reside in warm, arid and semi-arid soils of the Americas. The fungi exist in a dimorphic life cycle consisting of saprophytic mycelial phase in the soil, and a parasitic spherule phase when arthroconidia are inhaled by a mammalian host (Cole and Sun, 1985). Given proper conditions, slender filaments of saprophytic cells (hyphae) grow in the upper part of the soil (Kolivras, 2001). After a week or more of growth, many of the hyphal cells mature into rectangular arthrospores. The arthrospores alternate with smaller, sterile cells. When the soil dries, the alternating sterile cells breach easily, freeing the intervening arthrospores. The arthrospores range from 1.5 to 4.5 µm in width and 5.0 to 30 µm in length.

The parasitic stage usually initiates in the lungs, and it can spread to other parts of the body. If inhaled, the arthrospores can penetrate to the smallest bronchiole or pulmonary alveoli of the lung. In the parasitic phase, the arthrospore or arthroconidia develops into a spherical, double-walled cell called a spherule (sporangium). Spherules measure from 10 to 200 µm, and they typically contain a few to several hundred endospores, each 2 to 5 µm in diameter. The spherule eventually ruptures, discharging

the endospores into the neighboring tissue. Each endospore is potentially capable of blooming into a new spherule.

Exposure usually occurs following events that disrupt the soil, resulting in aerosolization of the fungal arthrospores (Schneider et. al., 1997). Both natural and anthropogenic or human soil disturbance can produce spore dispersion. The ecology of the pathogen remains obscure, and there is limited knowledge of the environmental antecedents of disease outbreaks. Detection of the fungus in the environment remains a critical challenge to modeling the source of disease (Barker et al., 2012).

1.2 Impacts of coccidioidomycosis

Most cases of *Coccidioides spp.* infections are self-resolved. Approximately 40% of infections are symptomatic, and they can result in severe complications such as influenza-like illnesses, community acquired pneumonia (CAP), lung cavities, and disseminated infections in the central nervous system, skin, bones, joints and other organs (Komatsu et al., 2003). About one percent of infected individuals experience serious, life threatening conditions such as meningitis and organ damage. On average, patients diagnosed with the disease suffer symptoms for six months. Risk factors for severe infections include race, age, and immunosuppression. The fungi also infect livestock, pets and wild animals.

Rates of coccidioidomycosis have increased epidemically in Arizona and other areas within the last two decades. The Arizona Department of Health Services (ADHS) reported a coccidioidomycosis incidence of 43 cases per 100,000 of population in 2001, representing an increase of 186% since 1995 (Komatsu et al. 2003). In 1998, 1,551 cases

were reported to the ADHS; 5,535 were reported in 2006, and over 10,000 were reported in 2009 (Komatsu et al. 2003; Sunenshine et al. 2007; Hector et al. 2011). It is the fourth most common disease reported to the Arizona Department of Health Services (Park et al. 2005), and Arizona accounts for 60 percent of reported cases in the nation (Center for Diseases Control and Prevention, 2004). Most (95%) of Arizona cases are in Maricopa, Pinal, and Pima Counties.

Coccidioidomycosis became a nationally reportable disease in 1995 at the southwest regional level, at which time a case definition was adopted that required laboratory confirmation. The reporting requirement is suggested to play a role, at least partially, in the increasing linear trend in exposure rates over the previous two decades (Sunenshine et al. 2007, Tamerius and Comrie, 2011).

The consequent public health burden to affected regions is considerable. Hospital charges for coccidioidomycosis in the U.S. exceeded $86 million in 2007. Recent studies noted that direct hospital charges for the treatment of the disease in Arizona was 26.8 million dollars with a median charge of $14,292 (Nguyen et al., 2013). It is an increasingly important health issue due to migration into the state from other regions and increased numbers of immuno-suppressed patients.

1.3 Environmental origins of *Coccidioides spp.*

Environmental factors are hypothesized to explain the occurrence and distribution of *Coccidioides spp.* at all points in the disease chain. Despite the significant increase in incidence of coccidioidomycosis, and the associated interest due to its economic and human costs, there is limited knowledge about the specific ecological niche required for

Coccidioides spp. to flourish, the environmental antecedents of disease outbreaks, and the precise mechanisms of spore aerosolization and subsequent dispersion (Galgiani, 1999).

Coccidioidomycosis cannot be transmitted from person to person, but it is acquired by inhalation of the arthrospores from the environment. Therefore it is reasonable to assume that the fungus should be easily isolated from endemic areas (Ajello et al. 1965; Lacy & Swatek 1974). However, only a few positive isolations from environmental samplings have been obtained in highly endemic areas in the United States (Stewart and Meyer 1932; Emmons 1942; Maddy 1965). The scarce environmental evidence for *Coccidioides spp.* seems to be in disagreement with the high incidence rates obtained for the disease. Detection of the fungus in the environment remains a critical challenge to modeling the source of disease (Barker et al., 2012).

On a regional scale, the major predictors of disease are climate, soil disturbance, and dust or wind events (Pappagianis 1994; Comrie 2005; Comrie and Glueck 2007, Tamerius and Comrie, 2011). However, at finer scales, the ecology of the fungus remains obscure and largely unstudied (Cox and Magee, 2004).

Infection usually occurs following activities or natural events that disrupt the soil, resulting in aerosolization of the fungal arthrospores (Schneider et. al., 1997). Wind erosion may contribute to spore dispersion. Evidence suggests that anthropogenic or human-induced soil disturbance generates dispersal of spores. Outbreaks of the disease have been associated with soil disruption, archeological digs, agriculture, and construction (Cairns et al., 2000; Park et al. 2005, Fisher et al., 2007). *Coccidioides spp.* is associated with alkaline soil that has a high salt content, rodents, rodent burrows and Amerindian middens (Swatek, 1970). Figure 1 illustrates a hypothetical synoptic model

for coccidioidomycosis, including ecological, exposure, and epidemiological components (Tabor, 2009).

1.4 Rodent hypothesis for coccidioidomycosis reservoir

Emmons (1942), a senior U.S. Public Health mycologist, proposed a rationale for a rodent reservoir of coccidioidomycosis. He based the rationale on the difficulty to isolate *Coccidioides spp.* from the soil. Further, it seemed improbable that a fungus which is so virulent for man and animals should have a natural habitat as a saprophyte in the soil. A search for an animal reservoir was indicated, and the presence of coccidioides spores in wind-blown soil would therefore be due to contamination from infected animals.

Emmons and Ashburn (1942) trapped and examined 303 wild rodents obtained from five sites at various seasons at the San Carlos Indian Reservation in Arizona. *Coccidioides spp.* were isolated as granulomatous lesions from the lung in 15 percent of the pocket mice and 17 percent of kangaroo rats. The frequency and chronic nature of the lesions suggested to the authors that rodents constitute a natural reservoir for the disease.

Egeberg and Ely (1956) collected over 500 soil samples in the southwestern sector of the San Joaquin Valley at ground surface and below ground surface, and from the walls of animal burrows. Overall, *Coccidioides spp.* were recovered in 7% of the total samples. However, *Coccidioides spp*. was recovered in 13.6% of samples from the walls of animal burrows, as opposed to 3.4% positive from the random samples (statistically significant). For the dry season, 4.2% of samples collected were positive, compared with

16% in wet season. They noted that the dry season positive samples came from subsurface samples, and wet season positives came from the surface samples. They suggested that the surface samples were sterilized by the heat from the dry season, and that animal burrows offered a conducive environment to the growth and concentration of *Coccidioides spp.* They postulated that burrows afforded protection from the very high summer surface temperatures, higher nitrogen content, and a collecting receptacle for fungi picked up from the surface by the tenant's fur.

Maddy (1957) noted that most of the fungi that cause deep mycoses in man and animals exist in nature as saprophytes in soil, on organic debris, or on vegetation. He and his co-workers isolated *Coccidioides spp.* from two species of rodents (*P. formosus* and *Citellus leucurus*) collected in the St. George area of Utah. He also reported that dogs can have disseminated disease, particularly with liver involvement. They occasionally shed the parasitic form of the organism in their feces. In a few days, the saprophytic form propagates sufficiently to produce infective arthrospores.

Maddy and Creceluis (1967) conducted studies on the role of the infected animal. Tissues and entire carcasses of mice, dogs, and cattle, infected with *Coccidioides spp.*, were buried at two sites in Arizona identified by soil sampling as *Coccidioides* free. Follow-up soil samplings, over a seven-year period, identified these sites as sometimes harboring *Coccidioides spp.*, while samplings from nearby areas were essentially negative. The research suggested that infected animal carcasses seed small surrounding areas of soil with the organism.

Among *Coccidioides* animal infection, wild rodents are the most reported in literature (Baptista-Rosas et al., 2012). Previous research found positive isolation in

several species of mice and rats. Of these, higher prevalence was found in kangaroo rat (*Dipodomys merriami*, 17 %) and desert pocket mice (*Chaetodipus penicillatus*, 15 %). Other species reported were ground squirrels (*Citellus ssp*) and grasshopper mice (*Onychomys torridus*).

1.5 Construction as a source of *Coccidioides spp*. dissemination

Point-source outbreaks of coccidioidomycosis are occasionally reported. They are associated with soil-disturbing activities in endemic areas such as archaeological digs, construction, military maneuvers, and outdoor group activities (Peterson et al., 2004). For example, an outbreak of coccidioidomycosis in a 12- person construction crew is reported in Cummings et al., 2010. Before infection, they participated in soil excavation for underground pipe installation on Camp Roberts Military Base, California. Ten workers developed symptoms of the disease or had serologically confirmed disease. In another example, Cairns et al. (2000) document a cluster of coccidioidomycosis in Washington state residents who had recently returned from Tecate, Mexico, where they assisted with construction projects including ground excavation for two swimming pools. Twenty-one serologically confirmed cases of the disease resulted, and *C. immitis* was subsequently isolated from soil samples in Tecate by use of the intraperitoneal mouse inoculation method.

Coccidioidomycosis is a recognized occupational illness. The Centers for Disease Control and Prevention considers workers engaged in soil-disrupting activities to be populations at risk for the disease (CDC, 2013). The highest occupational rates of coccidioidomycosis are observed among agricultural, construction and mining workers,

who are likely to have the greatest dust exposure. Human soil-disrupting activity, such as digging in endemic areas, is the most important determinant of coccidioidomycosis incidence in California (Das et al., 2012).

1.6 Spatial technology

A geographic information system (GIS), with its ability to manage and portray spatial data, has become the dominant tool in geography. A geographic information system (GIS) is an integrated set of tools and methodologies for collecting, storing, retrieving, analyzing, and displaying spatial as well as non-spatial attribute data. It has transformed a variety of health analyses and the structuring of public data (Meade and Earickson, 2000).

Perhaps the most important source for the development of the GIS, and its data, is remote sensing (Meade and Earickson, 2000). Satellites provide an enormous amount of digital data in multiple bands of the electromagnetic spectrum. Continuous and repetitious image collection allows differentiation of land cover or usage, and provides a platform for monitoring for change. Methodologies for the processing of remotely sensed data generate a variety of raster data products. Figure 2 illustrates use of remote sensing data for studies assessing the contribution of rodents to human disease.

1.7 Study area

The study is situated in the Sonoran Desert, an arid region stretching from southeastern Arizona, across Sonora, Mexico and through most of Baja California. The Sonoran Desert is characteristic for its semi-arid climate, mild winters and a bimodal rainfall pattern. Mountain ranges, volcanic hills, bajadas (coalesced alluvial fans), valley

floors, washes and arroyos (steep-sided gulches) are typical of its geomorphology (Hoffmeister, 1986). The visually dominant elements of the landscape are two characteristic life forms: legume trees and large columnar cacti (Arizona-Sonora Desert Museum, 2013).

The Lower Sonoran Life Zone (LSLZ), based on Merriam's elevation-precipitation life zones for the southwest (Merriam and Steineger, 1890), further delineates the study area. The LSLZ are the lower elevation parts of the desert (generally < 1000 m). The LSLZ stretches, in Pima County, Arizona, from the relatively pristine, sparsely-populated Organ Pipe Cactus National Monument, situated in the western half of the county, to the Tucson metropolitan area in the eastern half.

A ring of communities, unincorporated urban development, and undeveloped areas surround the city of Tucson. Here, thousands of acres of the Sonoran Desert are at the forefront of urbanization, and hundreds of acres have been bladed for the construction of houses and commercial strip malls (AIA, 2007). Between 1980 and 1990, the city's area increased by over 50 percent to approximately 600 km^2 through the annexation of unincorporated land. The regional population has also experienced a doubling in a similar time period. The population of Pima County is approximately one million (Census, 2012).

2. Data and methods

2.1 Rodent abundance maps

We built spatial data sets for abundance of ten nocturnal desert rodent species and several rodent groupings. The models were constructed with automated regression

methods using remotely-sensed terrain variables including thermal radiance, a vegetation index, image texture (or pixel diversity), and a suite of topographical variables (elevation, slope, aspect easting and northing, curvature and distance to the nearest wash). Images were spatially filtered with a 5x5 low pass median window in order to simulate sampling from a rodent trapping plot (approximately 90m or 8100m^2).

The thermal radiance, vegetation index and texture were generated from a June 1996 Landsat Thematic Mapper (TM) image. We looked for an image early in the study period for characterization of the landscape in advance of spectral changes associated with urban growth occurring in the region as the study progressed. The thermal radiance is derived from Landsat Band 6 (thermal infrared) calibrated digital numbers and rescaled to at-sensor spectral radiance using gain and offset coefficients provided in Chander et al. (2009).

We converted the June 1996 reflective image to atmospheric-corrected surface reflectance using the cosine approximation model (COST; Chavez, 1996). The COST model implements an improved dark-object atmospheric correction for Landsat TM multispectral data (bands 1-5 and 7). We prepared a Soil Adjusted Vegetation Index (SAVI; Huete, 1988) from the spatially-filtered annual median red and near-infrared images. The SAVI, derived from the Normalized Difference Vegetation Index (NDVI), incorporates a canopy adjustment factor to minimize soil noise inherent in the NDVI by accounting for differential red and near-infrared extinction through the canopy. We also generated a texture image from the SAVI using a variance, 2nd order operator on a 5x5 moving window.

Regression models for rodent abundance were obtained by comparisons to rodent trapping data from the Organ Pipe Cactus National Monument (ORPI) in southwestern Arizona. Pianalto and Yool (in progress) describe image processing and model development in more detail. Abundance values are based on count data for two night annual trapping events, and the values are adjusted for sprung traps and recaptures (Holm, 2006). Areas with predictor variables beyond the ranges of the original ORPI 30-plot data set were masked out to avoid unjustified extrapolation.

Models were extrapolated to from ORPI abundance and ORPI-area images (Landsat path 37, row 38) to eastern Pima County (Landsat path 36, row 38; NASA, 2011). Overlap of the two satellite paths allows comparisons of histograms for extrapolation (Figures 3 and 4). An additional mask was applied to remove small amounts of agriculture areas northwest and south of Tucson, and mining areas south of Tucson.

The regression model for desert pocket mouse abundance, *Chaetodipus penicillatus* (CHPE), is shown in Figure 5 as an illustration. Maps of rodent abundance distribution in eastern Pima County for the desert pocket mouse (CHPE), the Merriam's kangaroo rat, or *Dipodomys merriami* (DIME), the heteromyid rodent group (CHBA, CHIN, CHPE, PEAM, DIME, DISP), and the murid rodent group (NEAL, ONTO, and PEER, SIAR) are shown in Appendix 1 of this article.

2.2 Construction-related soil disturbance metric

We generated spatial data for construction-related soil disturbance using image differencing of Landsat TM band 5 (mid-infrared). Pianalto and Yool (2013) provide more detail on preparation of the construction-related soil disturbance data set. Thresholded annual change images were generated for the period of study, 1994-2006, and spatial coordinates and areal amounts of disturbance were prepared for comparisons with rodent and coccidioidomycosis data.

2.3 Coccidioidomycosis incidence data

Coccidioidomycosis incidence data are obtained from the Arizona Department of Health Services. Processing for the incidence data is described in more detail in Tamerius and Comrie (2011). Case data were processed to *date of exposure* based on symptom onset dates, diagnosis dates, and estimated coccidioidomycosis incubation period. Tamerius and Comrie further prepared detrended data to remove the strong linear trend observed across the study period for climatic studies; we used the non-detrended data since we hypothesized that construction-related soil disturbance contributes to the increasing linear trend. Monthly incidence per 100,000 were estimated using population estimates provided by the U.S. Census Bureau. The monthly averages were further aggregated to annual May-to-May averages for the period of this study to match the temporal period of the construction disturbance data set. County-wide data and data aggregated to zip-codes were prepared.

2.4 Data extraction

The rodent abundance and construction-related soil disturbance rasters were entered into a GIS for comparisons and extraction to county-wide and zip-code aggregations. The construction related soil disturbance metric is derived from the dimensions of the Landsat TM pixel, 30m or 900m^2. Figure 6 shows construction-related soil disturbance pixels, converted to points, with the Merriam's kangaroo rat abundance map for the southeastern Tucson peripheral area during the 1996-1995 change period. In this manner, the soil disturbance metric can be adjusted or weighted for rodent abundance.

An average rodent abundance value was extracted for each zip code. The count of soil disturbance pixels was also extracted for each zip code and each annual period; these were normalized by zip code area. Total disturbance counts, zip-code and county-wide, were also determined by summing up all years. Finally, each disturbance point was associated with a rodent abundance value for all species and groupings of rodents.

3. Results

3.1 Scale and extent

Spatial scale of the soil disturbance and rodent abundance models are defined by Landsat TM pixel size, 30m. Comparisons to coccidioidomycosis incidence, at zip code and county level aggregation, obligate up-scaling the predictors. Figure 7 shows Tucson-area zip codes derived from U.S. Census Bureau shapefiles. Temporally, the study period ranges 1994-2006. Incidence data, available at monthly granularity, is upscaled to annual May-to-May periods to match the scale of the construction disturbance change

images. This results in annual and 12-year combined study period comparisons at zip code and county levels of incidence with disturbance. The rodent abundant data are static, as the models are based on annual averages of abundance and environmental data for the period between 1993 and 2007. Figure 8 shows potential temporal and spatial scale variation for models.

3.2 Soil disturbance comparisons with incidence

Figure 9 plots a bar chart comparison of annual construction-related disturbance values and coccidioidomycosis incidence, both county-wide, for each of the twelve years of the study period (Model C). Spearman's rank order comparison of the two data sets shows a statistically significant correlation of 0.81 (p-value 6.8E-4). Figure 10 displays the same comparison estimated by linear regression. A statistical significant relationship is also estimated (R^2 = 0.49, p-value = 0.012), despite the small number of observations. Non-linear exponential and power models improve the correlation. Statistically significant linear and power correlations are preserved when two potential outliers, the last two annual periods, are removed (R^2 of 0.43 and 0.50 respectively, n = 10).

Spearman's ranked order correlation comparison between incidence and soil disturbance was performed for each individual zip code (12 annual data points each, Model D). Figure 11 displays the classified rank order coefficients for each zip code. The overall trend is for high correlation in zip codes at the periphery of Tucson, and neutral or no correlation at inner or core Tucson city zip codes. The highest correlation coefficient was 0.818 at zip code 85746 located southwest of Tucson.

Model C, or comparison of total disturbance and incidence over the 12 year study period with zip codes as observations, did not result in a statistically significant relationship. However, if the 38 zip codes are separated into two groups, a periphery group and a core group, the periphery demonstrates a statistically significant relationship (R^2 = 0.52; p-value 0.0012). Figure 12 provides a summary of statistical tests.

3.3 Rodent abundance comparisons with incidence

Average values of rodent abundance for rodent species and groupings were extracted for each zip code. This is the only comparison possible with the study design, as the rodent abundance maps represent 12-year averages. They are therefore static temporally. Furthermore, only periphery zip codes (n=17) were used in comparisons, based on the success of the soil-disturbance metric described in Section 3.1.2 with this subset.

The mean abundance for the desert pocket mouse, or *Chaetodipus penicillatus* (CHPE), performed best of the rodent abundance tests, with an adjusted R^2 of 0.25 (p-value 0.02). This model was derived from the SAVI vegetation index, aspect-northing, and thermal radiance from the scene (Figure 5). The valley type grouping (CHPE, PEAM, DIME, DISP, and ONTO) and all rodent species also generated models, but these were barely statistically significant (0.05 criteria).

Similar comparisons were performed with the individual biophysical variable images used to build the rodent models. The texture image derived from the SAVI and the curvature derived from the digital elevation model generated passing models with adjusted R^2's of 0.26 and 0.18 respectively (p-values 0.02 and 0.05). Combining the

texture and curvature into a multivariate model improved the correlation to an adjusted R^2 of 0.40 (p-value 0.01).

3.4 Combined abundance and disturbance comparisons with incidence

We performed an automated stepwise regression on the periphery zip code data set (n=17) of the 12 year, total study period disturbance count values and the average abundance values in the zip code for each rodent species and grouping. Default entrance and exit tolerances for the *F*-statistic *p*-values (0.05 and 0.10, respectively) were used. With all the possible combinations of disturbance and rodent abundance variables, the automated procedure resulted in a model with the construction disturbance count as a single variable; i.e. no rodent abundance improved the disturbance model. Combining the best performing rodent abundance model, CHPE, with soil disturbance count, for the periphery zip code, full study period data set, slightly improved statistics above disturbance counts alone (Adjusted R^2 of 0.52 and 0.48, respectively).

4. Discussion

Spatial and temporal resolution are key components in this study's design. The resolution of the Landsat TM (30m pixel) is well-matched for the evaluation rodent plots (90m) and construction disturbance (average 480m dimension). The extent of Landsat TM imagery (approximately 185km scenes) is appropriate for data storage and processing requirements of a regional study.

Rodent abundance comparisons, based on zip code averages, for the most part did not indicate any statistical relationship with coccidioidomycosis incidence. The

rodent model was static temporally. This perhaps constrained the investigation, as rodent populations vary significantly seasonally and inter-annually. Whether or not improved temporal resolution would have affected results is untested. Models built with applying a rodent abundance weighting to each construction soil disturbance pixel have yet to be tested.

Coccidioidomycosis incidence data is available at zip code granularity; this may be a sufficient match considering exposure models for spore dissemination from a disturbance site. The pathogen can be carried or dispersed by wind, or by movement of equipment, vehicles, and people. However, the zip codes in Pima County vary widely in size, with larger areas in the periphery of Tucson. Although the disturbance count was normalized by area, the large discrepancy in size between the central and periphery zip codes may introduce scale issues in modeling.

Statistically significant relationships were observed between coccidioidomycosis incidence and soil disturbance at both spatial scales; this enhances the conclusion that construction related soil disturbance, estimated by the metric reported here, plays a direct relationship in dissemination of the disease. Variations in spatial scale, in this case from county level to zip code level, provided a useful means to ascertain relations of coccidioidomycosis incidence and soil disturbance.

Coccidioidomycosis incidence in Arizona has seen a progressive, linear increase over the previous decade and a half. This increase is often attributed to the disease reporting mandate. However, the region has grown significantly in the same period, despite periodic booms and busts. Results from this study suggest that construction soil disturbance is a factor in this trend.

In summary, using satellite imagery, or a synoptic view, appears to sufficient for estimating soil disturbance resulting from construction projects. However, it may be insufficient for assessing the hypothesis of a rodent contribution to coccidioidomycosis incidence. In this case, improved theoretical development into the precise reason(s) for a rodent association, and/or more mechanistic models, are expected to improve the case for a remote sensing role.

5. Acknowledgements

Sincere appreciation is expressed to Dr. Andrew Comrie and research group at the Department of Geography and Development, University of Arizona, for the monumental exercise of processing coccidioidomycosis case data into county and zip code incidence data. Appreciation is also expressed to Dr. Daoquin Tong at the Department of Geography and Development, University of Arizona for her assistance with statistical aspects of the models.

6. Figures and Tables

Figure 1: Synoptic model for coccidioidomycosis: ecology, exposure and epidemiology (Tabor, 2009). Hypothesized rodent and construction soil contributions to the pathway are outlined in circles or squares.

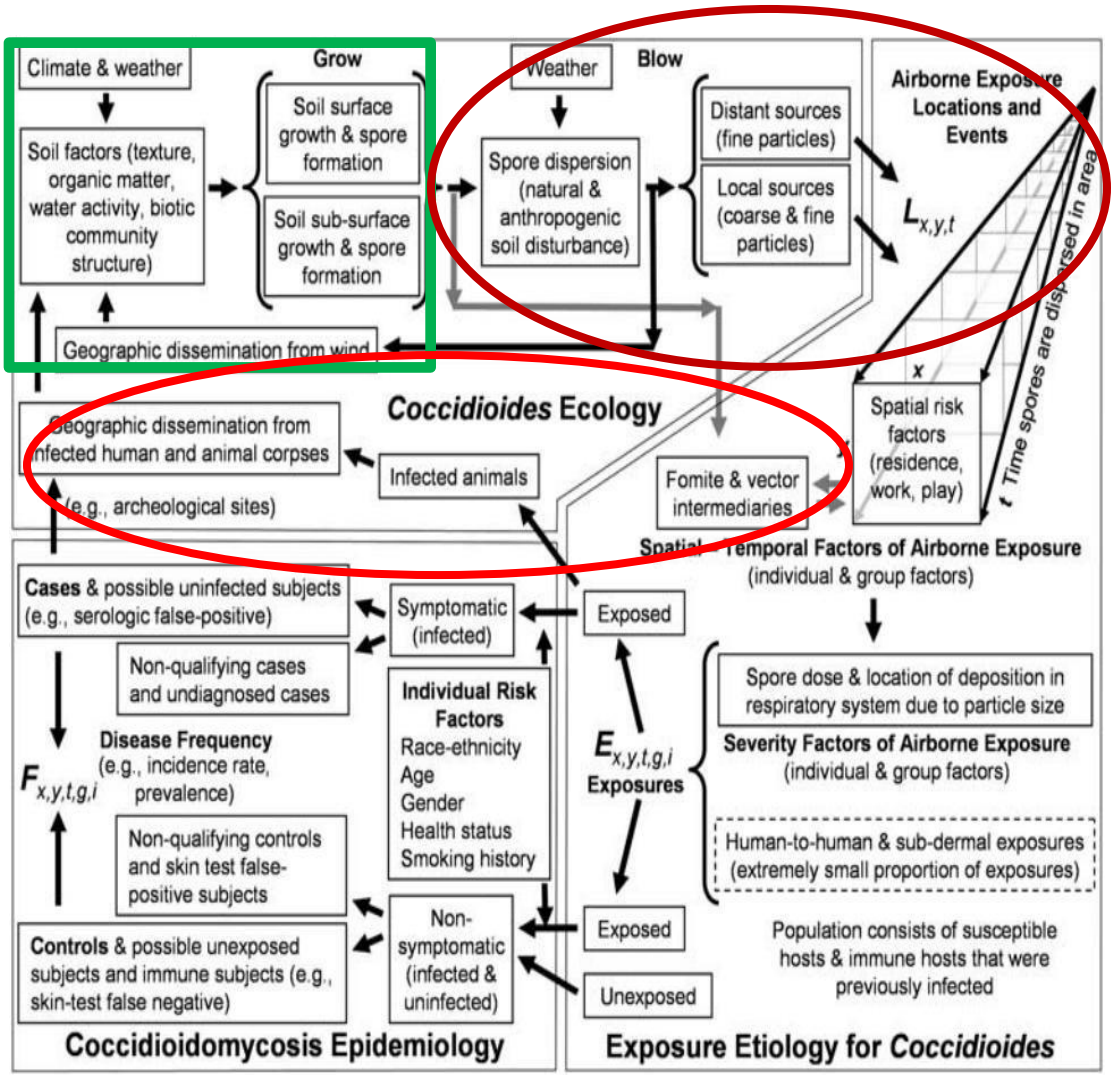


Figure 2: Role of remote sensing in assessing rodent contributions to human disease (Mills and Childs, 1998).

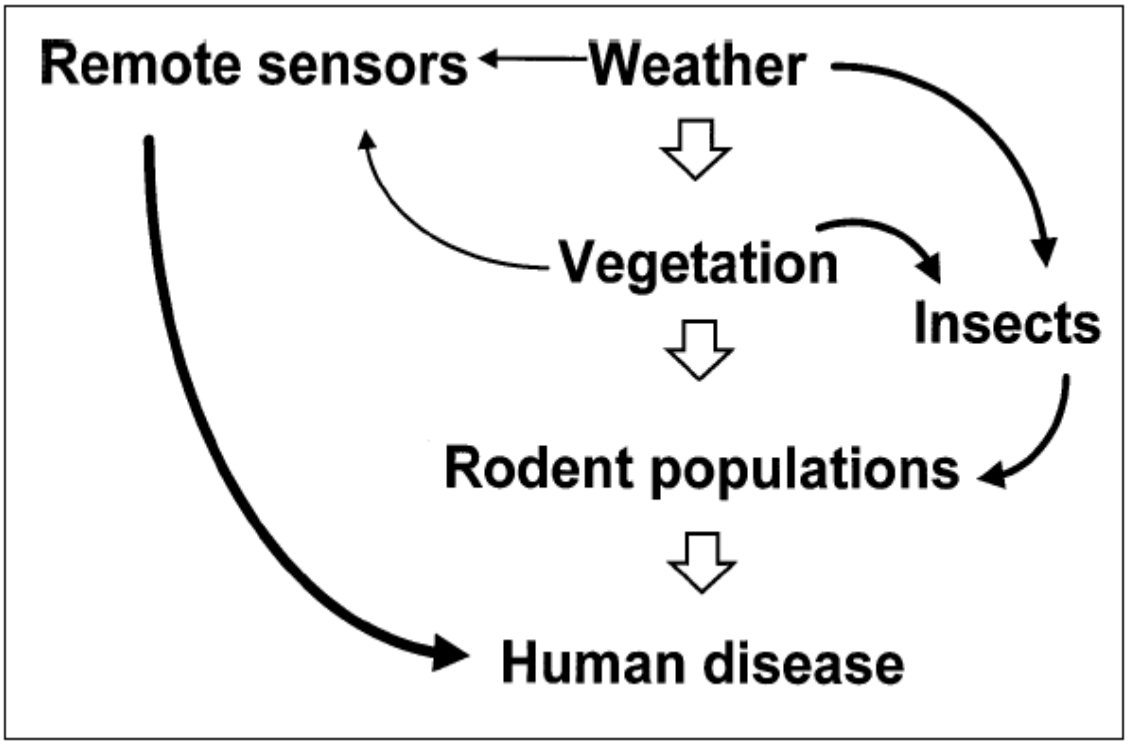


Figure 3: Landsat image overlap. Rodent abundance models are generated with biophysical variables derived from Landsat TM 5 path 37 and row 38 imagse and compared with abundance data obtained from the the Organ Pipe National Monument. Models are extrapolated to eastern Pima County and the Tucson metropolitan area, which are situated in Landsat path 36 and row 38. An area of overlap exists between both image paths. Surface radiance images derived from Landsat band 6 (thermal infrared) for both paths are shown below. The background of the path 36 image obscures the eastern portion of the image on the path 37 image in this display. A region of overlap is highlighted by the box on the images.

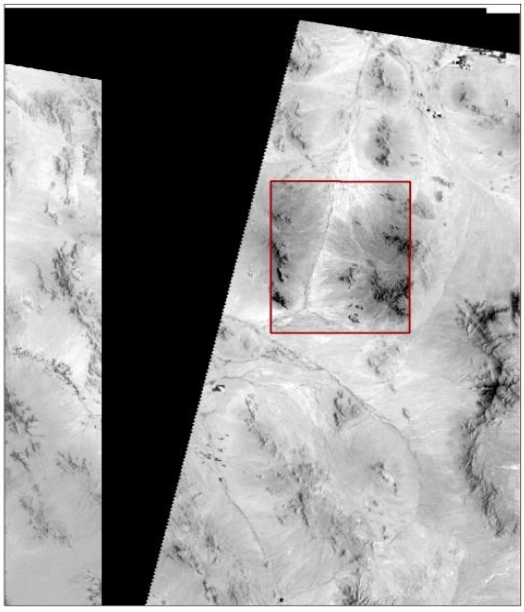


Figure 4: Histograms of the highlighted overlap areas of the two Landsat path images shown in Figure 5. The histograms show a strong degree of correlation, and supports extrapolation from the ORPI scene to the eastern Pima County scene in a different Landsat image collection path.

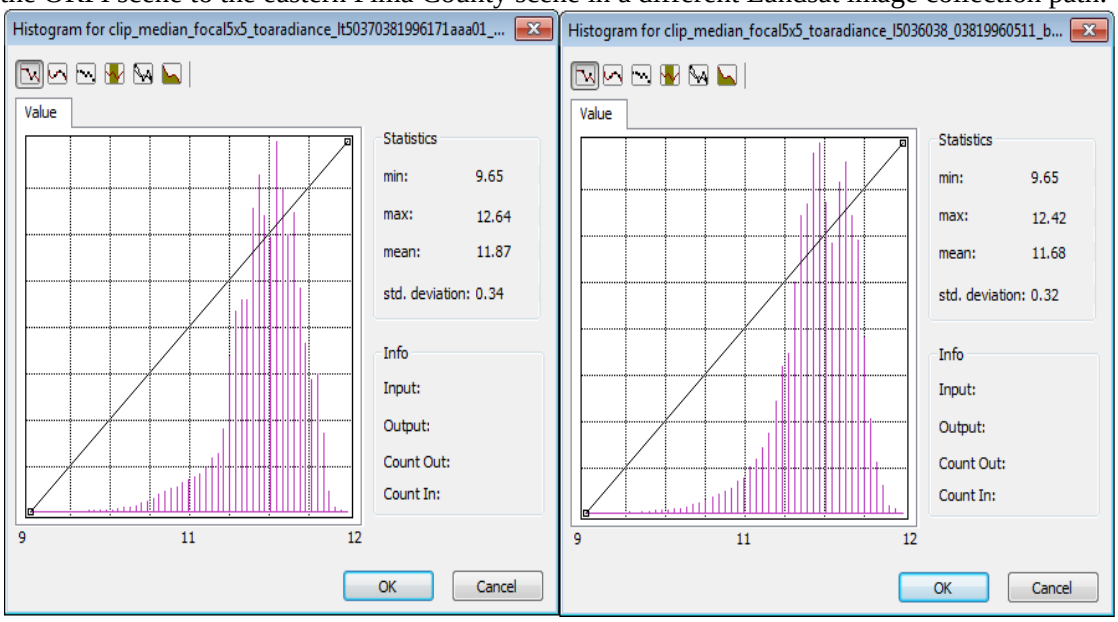


Figure 5: Abundance model predicted by stepwise linear regression for *Chaetodipus penicillatus* (CHPE, desert pocket mouse) from trapping data and images at the Organ Pipe Cactus National Monument in southwestern Arizona. The overall adjusted coefficient of determination is 0.49 (p-value 4.7E-4). SAVI is the Soil-Adjusted-Vegetation Index. Aspect is the northing ranging form -1 due south to +1 due north, and Thermal Radiance is derived from thermal infrared imagery.

Variable	Coefficient	Standard Error	p-value
SAVI	708.5	213.0	0.0026
Aspect_N	-24.6	6.5	0.00077
Thermal Radiance	12.3	7.1	0.092
Constant	-204.8	93.0	0.037

Figure 6: Construction-related soil disturbance pixels, shown as blue/red or dark points are displayed on an abundance map for Merriam's kangaroo rat (*Dipodomys merriami*, or DIME). The soil disturbance map is generated from thresholded Landsat TM band 5 (mid-infrared) imagery, and the DIME abundance map is generated from Landsat TM band 6 (thermal infrared) imagery. GIS allows a platform to extract a rodent abundance value and weight each disturbance point by its abundance value to generate coccidioidomycosis predictive models using a combination of both environmental variables.

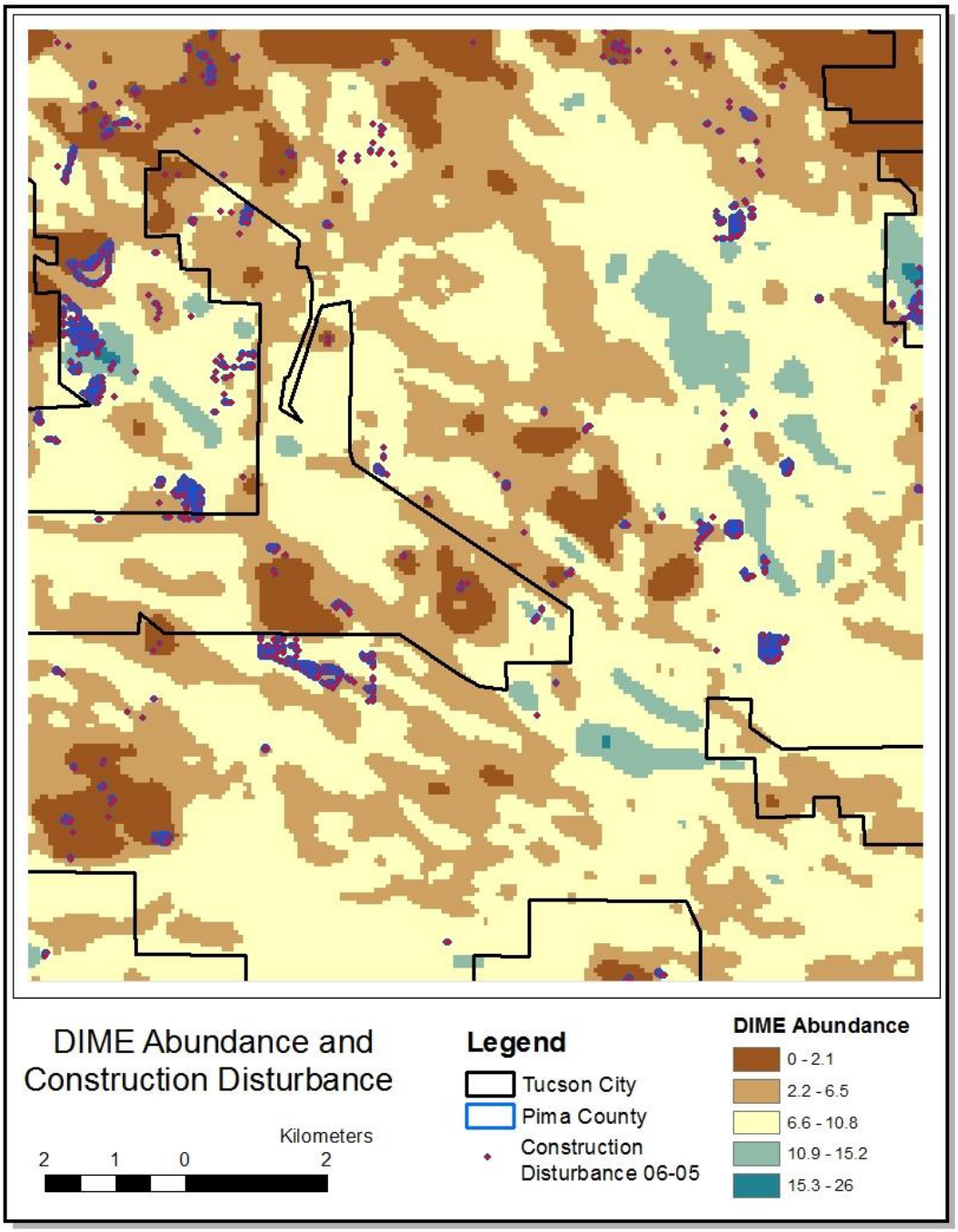


Figure 7: Study area including eastern Pima County, Arizona and the metropolitan region of Tucson. Aggregation of data to eastern county-wide and to zip codes is accomplished for statistical investigations. Rodent abundance and construction soil disturbance data is derived from Landsat TM 5 satellite imagery, and a typical northern edge of the image is shown.

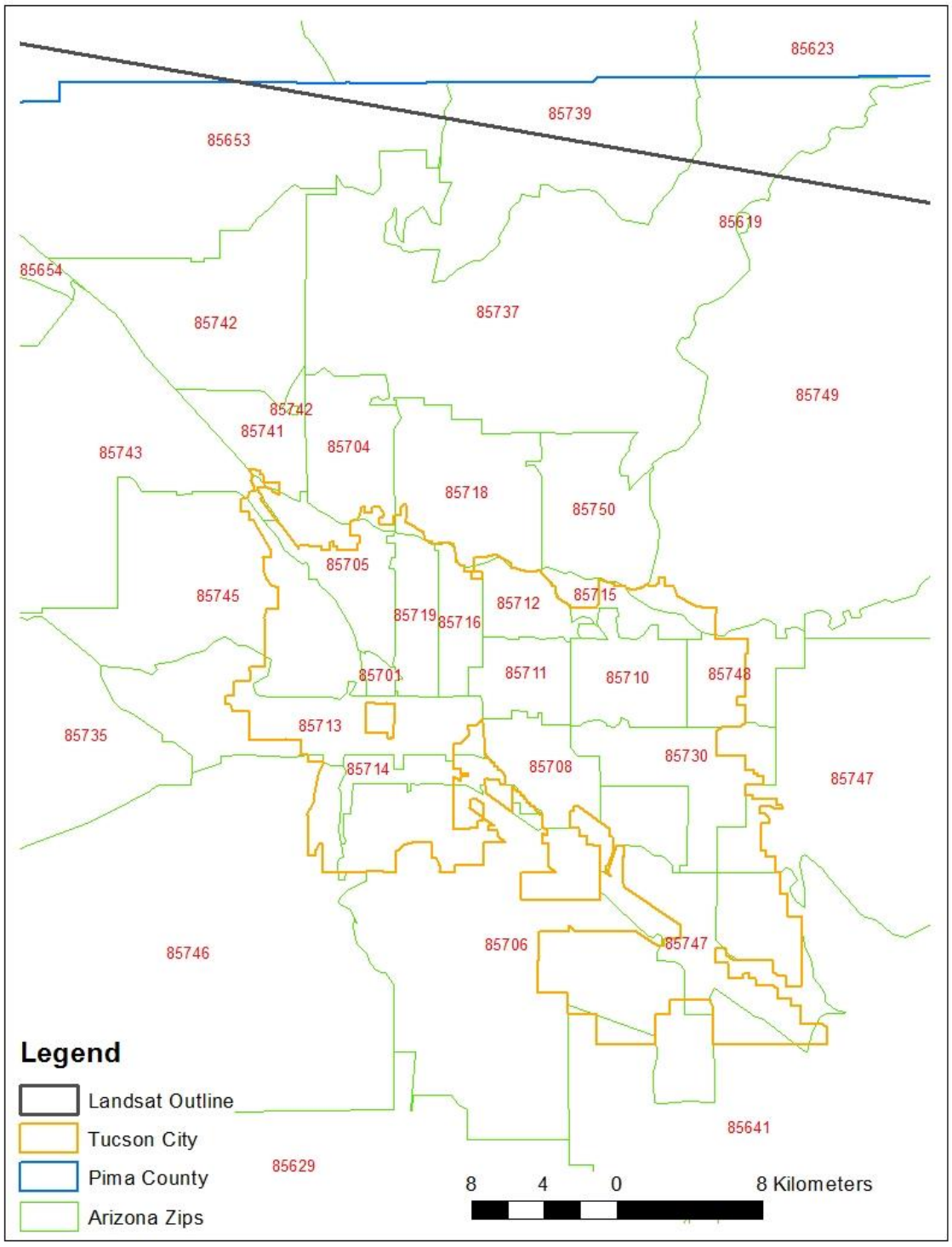


Figure 8: Temporal and spatial aggregations of the data sets in multiple ways allow for interscale comparisons.

Temporal/Spatial	County-wide (1 county)	Zip Code (38 zips)
Entire Study Period (1 period)	**Model A:** 1 data point	**Model B:** 38 data points
Annual (12 years)	**Model C:** 12 data points	**Model D:** 456 data points Or 38 x 12 data point studies

Figure 9: Comparison of coccidioidomycosis incidence and construction-related soil disturbance area, both aggregated to annual May to May periods over the period of study (1994-2006). The relationship is statistically significant, with a Spearman's rank order *rho* of 0.81 (p-value 6.8E-4).

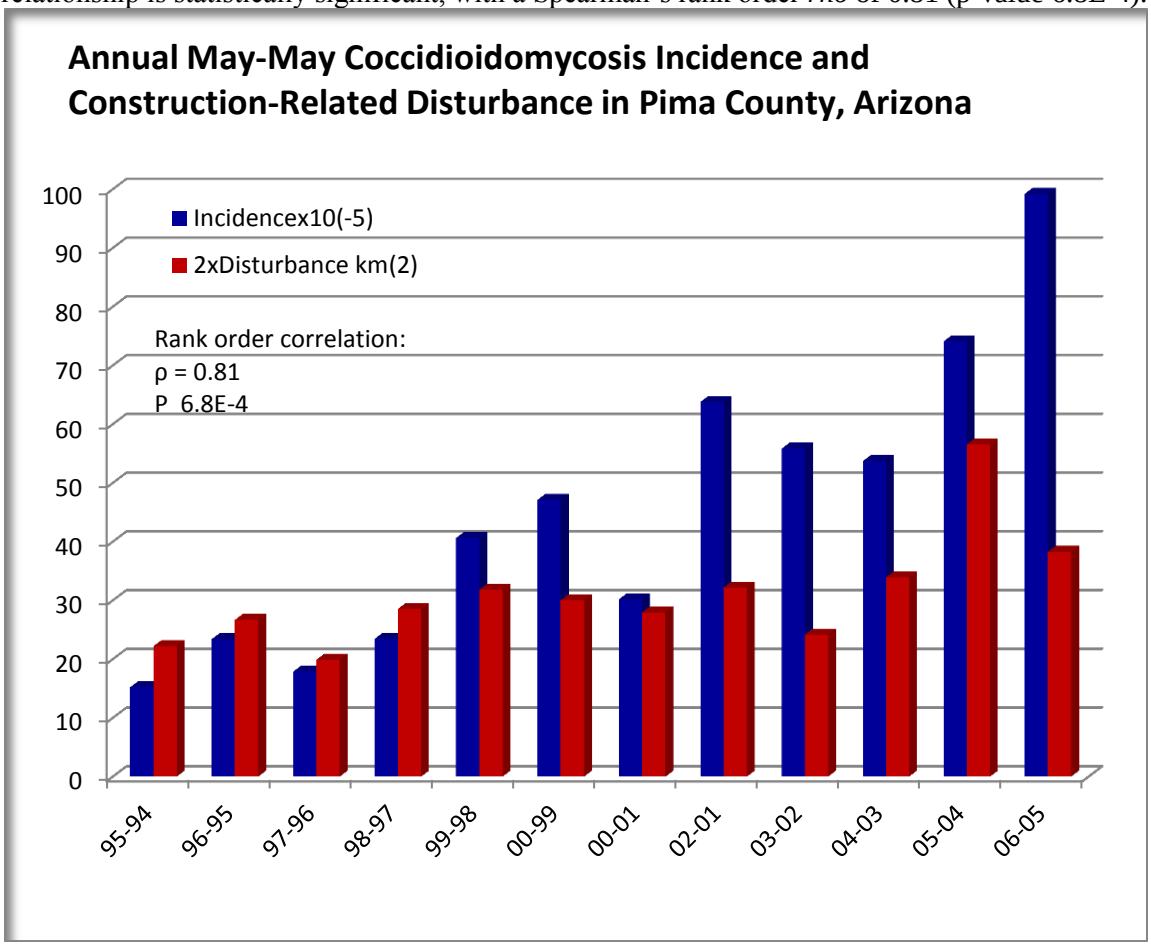


Figure 10: Comparison of coccidioidomycosis incidence and construction-related soil disturbance area, both aggregated to annual May to May periods over the period of study (1994-2006). This is the same comparison as displayed in Figure 9, but with a linear regression test. The relationship is statistically significant, with an adjusted coefficient of determination of 0.49 (p-value 0.012). A nonlinear power regression results in improved correlation.

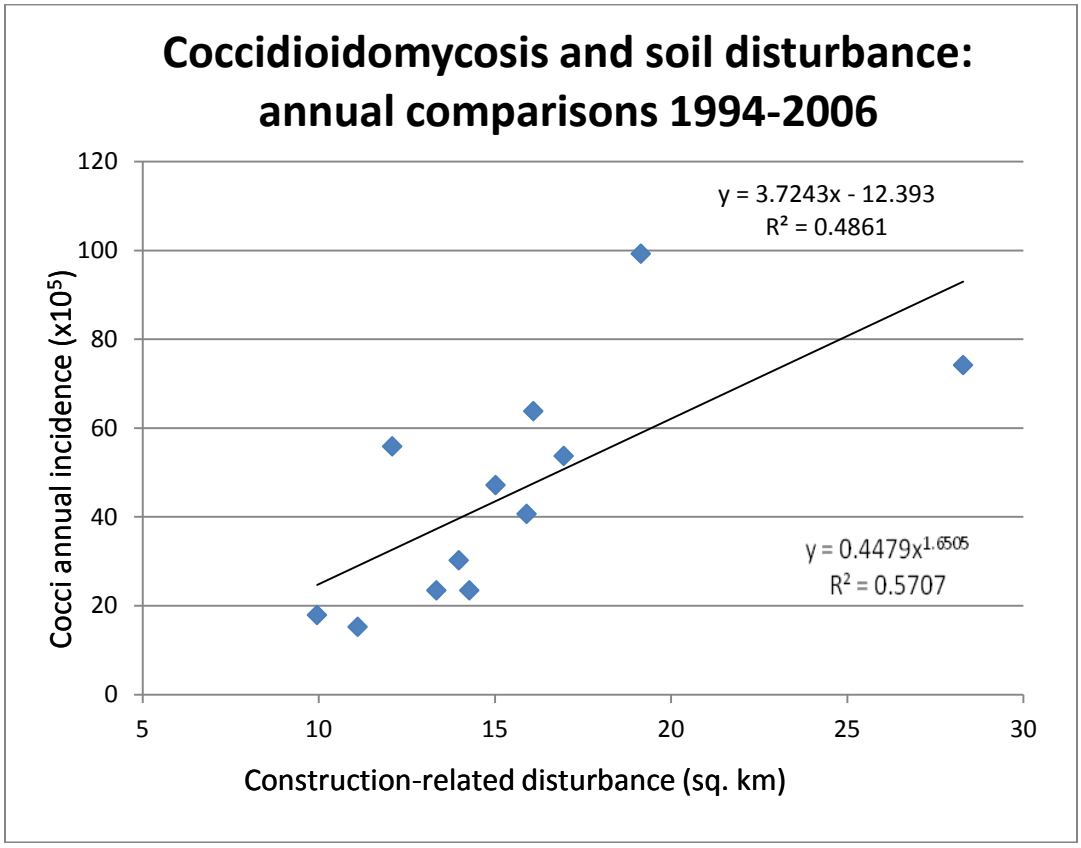


Figure 11: Independent Spearman's rank order correlation testing is performed on each zip code, using its respective coccidioidomycosis incidence and soil disturbance values, both aggregated to 12 annual observation data points. Higher rank order correlation, shown in red and orange, is observed in the periphery areas of the Tucson metropolitan area. These areas are also experiencing the most change associated with urban growth, which is expected to generate more construction-related fugitive dust. The core or central city zip codes, generally show little or no correlation.

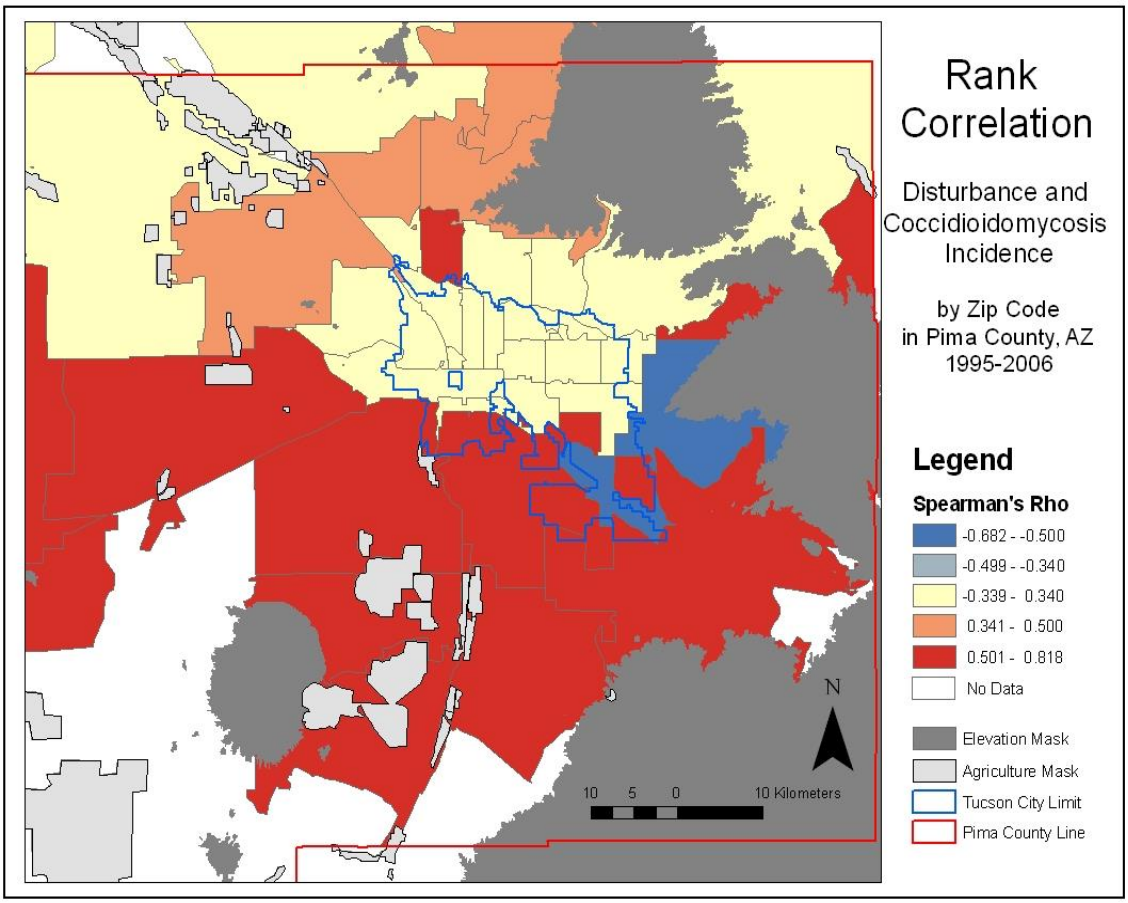
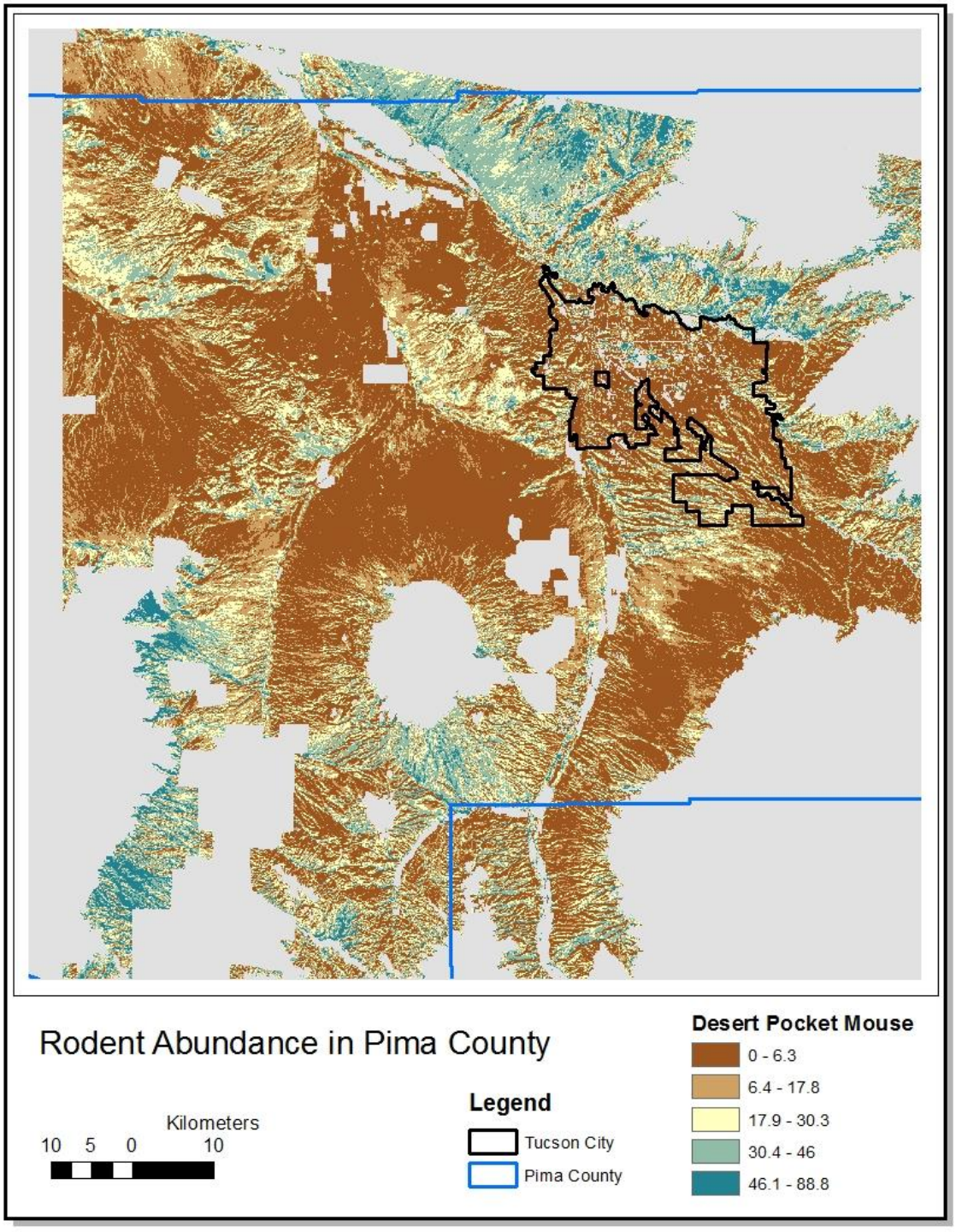


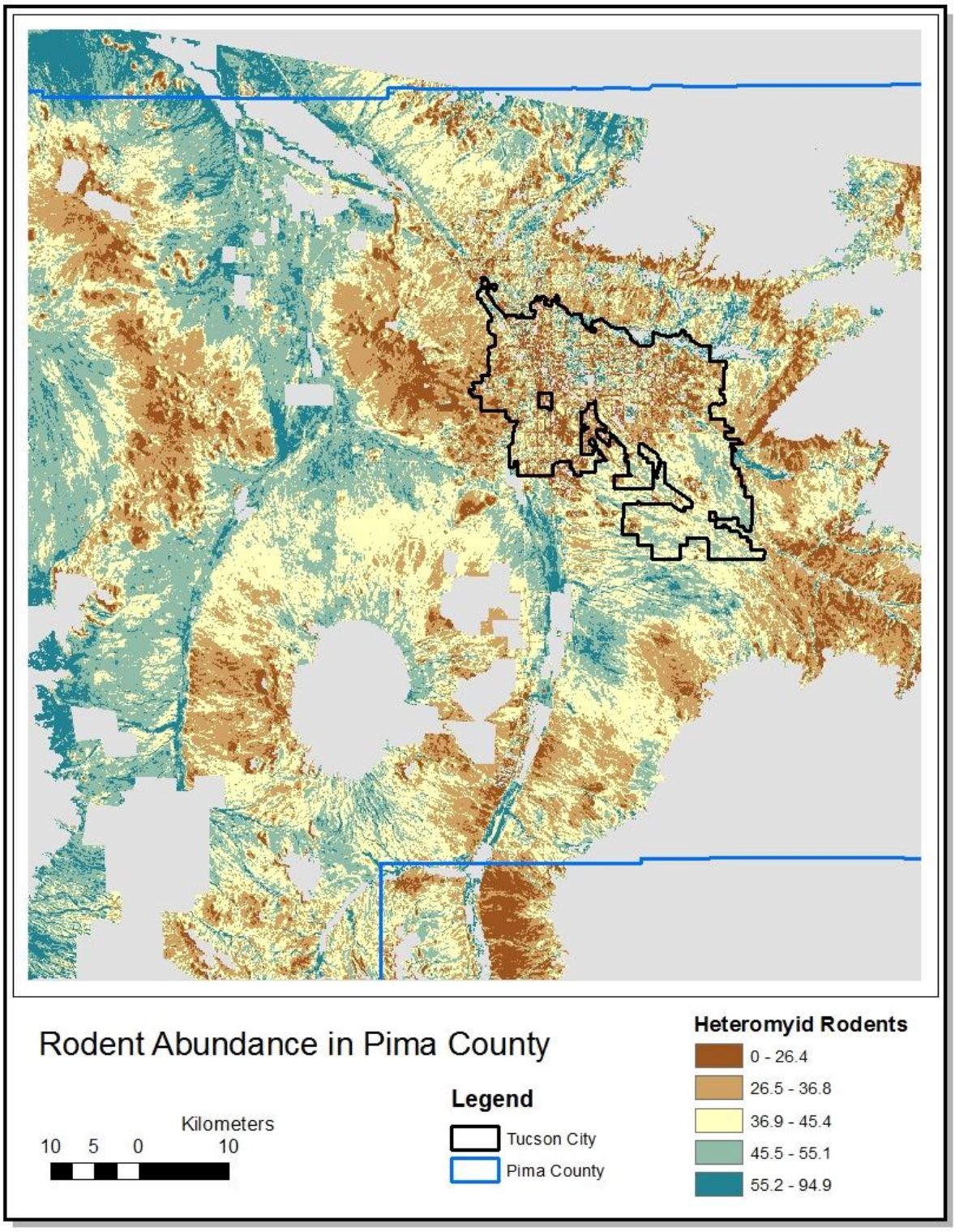
Figure 12: summary of linear and multivariate linear statistical comparisons between coccidioidomycosis with rodent abundance and soil disturbance metrics. Statistically insignificant models are shown with a dash.

Predictor Variables	Spatial Aggregation	Temporal Aggregation	Observations	Adj. R^2	p-value
soil disturbance	county	annual	12	0.49	0.012
soil disturbance + precipitation	county	annual	12	0.65	0.004
soil disturbance	zips, all	annual	456	-	-
soil disturbance	zip, all	total 12 year period	38	-	-
soil disturbance	zip, periphery	total 12 year period	17	0.48	0.001
CHPE, desert pocket mouse	zip, periphery	total 12 year period	17	0.25	0.02
soil disturbance + CHPE	zip, periphery	total 12 year period	17	0.52	0.002
texture of SAVI	zip, periphery	total 12 year period	17	0.26	0.02
texture_SAVI + curvature	zip, periphery	total 12 year period	17	0.40	0.01

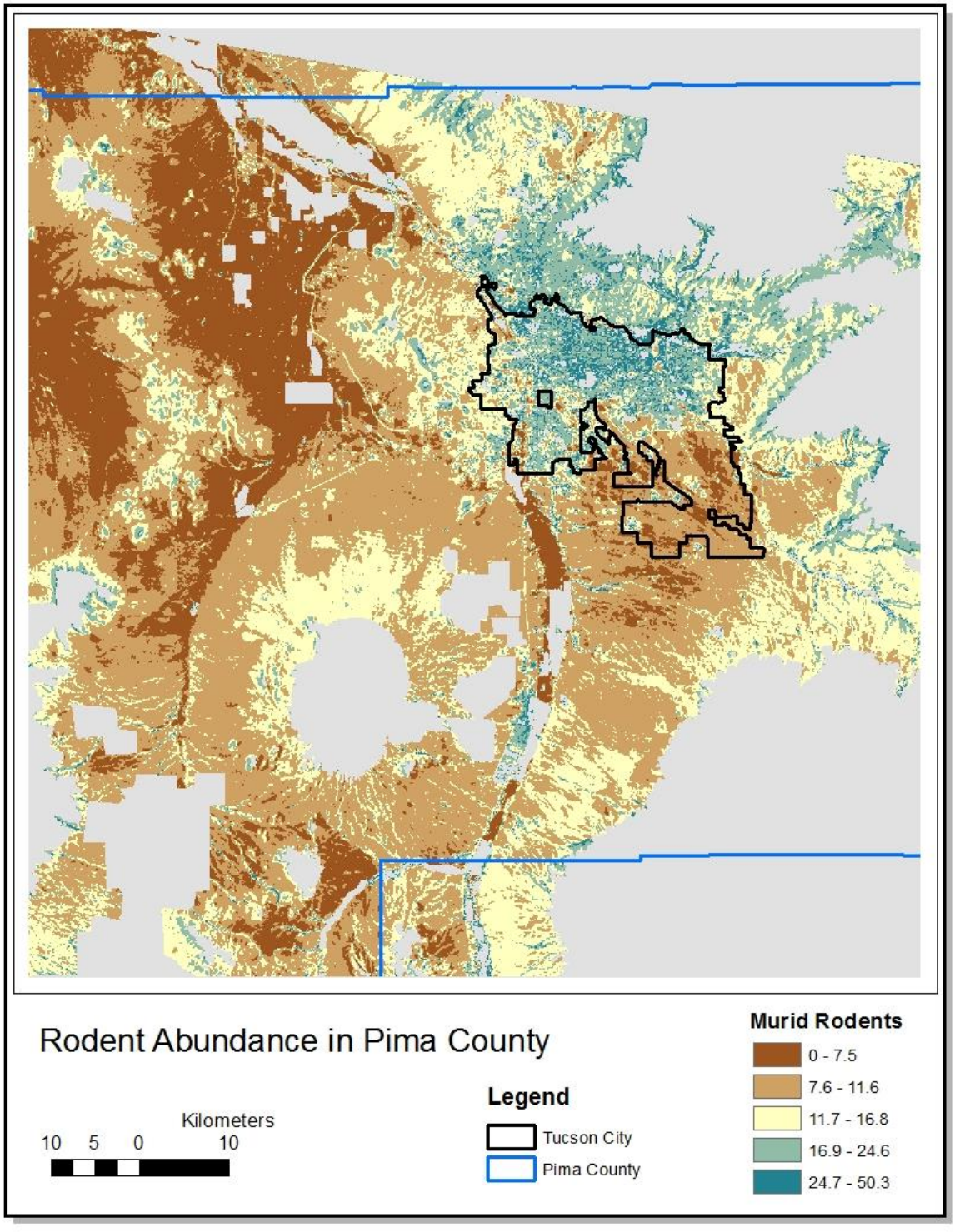
Appendix 1.a Abundance model for the desert pocket mouse (*Chaetodipus penicillatus*, or CHPE). This model is generated with SAVI (Soil-Adjusted-Vegetation Index), Aspect northing (ranging form -1 due south to +1 due north), and Thermal Radiance (derived from thermal infrared imagery). Masked areas, shown in gray, are areas outside the ranges of predictive variable in the original model, or an agriculture and mining area mask.



Appendix 1.b: Abundance map for heteromyid rodent grouping, including CHBA, CHIN, CHPE, PEAM, DIME, DISP species of rodents. These are the pocket mice and kangaroo rats. This model is generated with surface radiance and a soil adjusted vegetation index (SAVI).



Appendix 1.c: Abundance map for murid rodent grouping, including NEAL, ONTO, PEER, SIAR species of rodents (pack rat, grasshopper mouse, cactus mouse and cotton rat). This model is generated with surface radiance and a texture of the soil adjusted vegetation index (Text-SAVI).



7. References

AIA, 2007, Tucson/Pima County, Arizona: One Million Reasons to Plan for Sustainable Growth, Southern Arizona Chapter of the American Institute of Architects, http://www.aia.org/aiaucmp/groups/aia/documents/pdf/aias078080.pdf.

Ajello, L., Maddy, K., Crecelius, G., Hugenholtz, P.G., and L.B. Hall, 1965. Recovery of Coccidioides immitis from the air, Journal of the International Society for Human and Animal Mycology, 4(2): 92-95.

Arizona-Sonora Desert Museum, 2000. A natural history of the Sonoran Desert, Phillips, S.J. and P.W. Comus, Eds., University of California Press, Berkeley CA, 650 pp.

Baptista-Rosas, R.C., Catalan-Dibene, J., Romero-Olivares, A.L., Hinojosa, A., Cavazos, T. and M. Riquelme, 2012. Molecular detection of Coccidioides spp. from environmental samples in Baja California: linking Valley Fever to soil and climate conditions, Fungal Ecology 5: 177-190.

Barker, B.M., Tabor, J.A., Shubitz, L.F., Perrill, R., and M.J. Orbach, 2012. Detection and phylogenetic analysis of Coccidioides posadasii in Arizona soil samples, Fungal Ecology 5: 163-176.

Cairns, L., Blythe, D., Kao, A., Pappagianis, D., Kaufman, L., Kobayashi, J. and R. Hajjeh, 2000. Outbreak of coccidioidomycosis in Washington state residents returning from Mexico, Clinical Infectious Diseases, 30: 61-64.

CDC, 2013. Centers for Disease Control and Prevention. Coccidioidomycosis (Valley Fever), Fungal Homepage, http://www.cdc.gov/fungal/coccidioidomycosis/

Census, 2012, State & County QuickFacts, Pima County, Arizona, U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/04/04019.html.

Chander, G., Markham, B. L., and D. L. Helder, 2009. Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors, Remote Sensing of Environment, 113: 893-903.

Chavez, P. S., Jr., 1996. Image-Based Atmospheric Corrections - Revisited and Improved, Photogrammetric Engineering & Remote Sensing, 62(9): 1025-1036.

Cole, G.T. and S.H. Sun, 1985. Arthroconidium - spherule - endospore transformation in Coccidioides immitis. In: Fungal dimorphism, Szaniszlo P.J. (Ed), Plenum Publishing Corporation, New York City 1985. p.281.

Comrie, A.C., 2005. Climate factors influencing coccidioidomycosis seasonality and outbreaks. Environmental Health Perspectives, 113: 688–692.

Comrie A.C. and M.F. Glueck, 2007. Assessment of climate-coccidioidomycosis model: model sensitivity for assessing climatologic effects on the risk of acquiring coccidioidomycosis. Annals of the New York Academy of Sciences 1111: 83–95. doi: 10.1196/annals.1406.024.

Cox, R.A. and D.M. Magee, 2004. Coccidioidomycosis: host response and vaccine development. Clinical Microbiology Reviews, 17:804–839.

Cummings, K.C., McDowell, A., Wheeler, C., McNary, J., Das, R., Vugia, D.J., and J.C. Mohle-Boetani, 2010. Point-source outbreak of coccidioidomycosis in construction workers, Epidemiology and Infection Apr 138(4): 507-511.

Das, R., McNary, J., Fitzsimmons, K., Dobraca, D., Cummings, K., Mohle-Boetani, J., Wheeler, C., McDowell, A., Iossifova, Y., Bailey, R., Kreiss, K., and B. Materna, 2102. Occupational coccidioidomycosis in California, outbreak investigation, respirator recommendations and surveillance findings, Journal of Occupational and Environmental Medicine, 54(5), p. 564.

Egeberg, R.O. and A.F. Ely, 1956. Coccicioides immitis in the soil of southern San Joaquin Valley, The American Journal of the Medical Sciences, 231: 151-154.

Emmons, C.W., 1942. Isolation of Coccidioides from soil and rodents. Public Health Reports 57(4): 109-111.

Emmons, C.W. and L.L. Ashburn, 1942. The isolation of haplosporangium parvum n. sp. And Coccidioides immitis from wild rodents. Their relationship to coccidioidomycosis. Public Health Reports (1896-1970), 57(46):1715-1727

Fisher, M.C., Koenig, G.L., White, T.J., and J.T. Taylor, 2002. Molecular and phenotypic description of Coccidioides posadasii sp. nov., previously recognized as the non-California population of Coccidioides immitis. Mycologia 94: 73–84.

Fisher, F.S., Bultman, M.W., Johnson, S.M., Pappagianis, D., and E. Zaborsky, 2007. Coccidioides niches and habitat parameters in the southwestern United States. Annals of the New York Academy of Sciences, 1111: 47–72.

Galgiani, J.N., 1999. Coccidioidomycosis: a regional disease of national importance, Rethinking approaches for control, Annals of Internal Medicine, 130(4): 293-300.

Hector, R.F., Rutherford, G.W., Tsang, C.A., Erhart, L.M., McCotter, O., Komatsu, K., Anderson, S.M., Tabnak, F., Vugia, D.J., Yang, Y., and J.N. Galgiani, 2011. Public health impact of coccidioidomycosis in California and Arizona. International Journal of Environmental Research and Public Health, 8: 1150–1173.

Hoffmeister, D.F., 1986. Mammals of Arizona, The University of Arizona Press, Tucson AZ, 600 p.

Holm, P., 2006. Nocturnal Rodents, Chapter 10, Organ Pipe Cactus National Monument Ecological Monitoring Report, 1997-2005. December 2006. National Park Service, U.S. Department of the Interior, Organ Pipe National Monument, Arizona http://www.nps.gov/orpi/naturescience/orpi-ecological-monitoring-report.htm.

Huete, A. R. 1988. A Soil-adjusted Vegetation Index (SAVI), Remote Sensing of the Environment, 25:295-309.

Kolivras, K.N., Johnson, P.S., Comrie, A.C., and S.R. Yool, 2001. Environmental variability and coccidioidomycosis (valley fever), Aerobiologia 17: 31-42.

Komatsu, K., Vaz, V., McRill, C., Colman, T., Comrie, A., Sigel, K., Clark, T., Phelan, M., Hajjeh, R. and B. Park, 2003. Increase in coccidioidomycosis – Arizona, 1998-2001, MMWR 52(6): 109-112.

Lacy, G.H. and F.E. Swatek, 1974. Soil ecology of Coccidioides immitis at Amerindian middens in California. Applied Microbiology. 27:379–388.

Maddy, K., 1957. Ecological Factors of the Geographic Distribution of Coccidioides immitis, Journal of the American Veterinary Medical Association, 130: 475-476.

Maddy, K,T., 1965. Observations on Coccidioides immitis found growing naturally in soil. Arizona Medicine 22: 281–288.

Maddy, K.T. and H.G. Crecelius, 1967. Establishment of Coccidioides immitis in negative soil following burial of infected animals and animal tissues, in Ajello, L. (ed.) Proceedings of Second Coccidioidomycosis Symposium, Tucson AZ: University of Arizona Press, pp. 309-312.

Meade, M.S. and R.J. Earickson, 2000. Medical geography, Second Edition., The Guilford Press, New York, NY, p. 14.

Merriam, C.H. and L. Steineger, 1890. Results of a biological survey of the San Francisco mountain region and the desert of the Little Colorado, Arizona. North American Fauna Report 3. U.S. Department of Agriculture, Division of Ornithology and Mammalia, Washington, D.C., 136 pp.

Mills, J.N. and J.E. Childs, 1998. Ecological studies of rodent reservoirs: their relevance for human health. Emerging Infectious Diseases 4(4): 529-537.

NASA, 2011. The Thematic Mapper, National Aeronautics and Space Administration, Washington D.C. http://landsat.gsfc.nasa.gov/about/tm.html.

Nguyen, C., Barker, B.M., Hoover, S., Nix, D.E., Ampel, N.M., Frelinger, J.A., Orbach, M.J., and J.N. Galgiani, 2013. Recent advances in our understanding of the

environmental, epidemiological, immunological, and clinical dimensions of coccidioidomycosis. Clinical Microbiology Review 26(3): 505-525.

Pappagianis D., 1994. Marked increase in cases of coccidioidomycosis in California: 1991, 1992, and 1993. Clinical Infectious Disease, 19:S14 –S18.

Park, B.J., Sigel, K., Vaz, V., Komatsu, K., McRill, C., Phelan, M., Colman, T., Comrie, A.C., Warnock, D.W., Galgiani, J.N., and R.A. Hajjeh, 2005. An epidemic of coccidioidomycosis in Arizona associated with climatic changes, 1998-2001. The Journal of Infectious Diseases, 191: 1981-1987.

Peterson L.R., Marshall, S.L., Barton-Dickson, C., Hajjeh, R.A., et al. 2004. Coccidioidomycosis among workers at an archeological site, northeastern Utah. Emerging Infectious Diseases 10: 637–642.

Pianalto, F.S., and S.R. Yool, 2013. Monitoring fugitive dust emission sources arising from construction: a remote-sensing approach, GIScience & Remote Sensing, 50(3): 251-270,

Pianalto, F.S. and S.R. Yool, in progress, Nocturnal desert rodent abundance distribution in southwestern Arizona estimated by regression methods with remote sensing-derived data.

Schneider, E., Hajjeh, R.A., Spiegel, R.A., Jibson, R.W., Harp, E.L., Marshall, G.A., et al., 1997. A coccidioidomycosis outbreak following the Northridge, Calif, earthquake, The Journal of the American Medical Association, 277(11): 904-908.

Stewart, R.A. and K.F. Meyer, 1932. Isolation of Coccidioides immitis (Stiles) from soil. Proceedings of the Society for Experimental Biology and Medicine, 29: 937-938.

Sunenshine, R.H., Anderson, S., Erhart, L., Vossbrink, A., Kelly, P.C., et al., 2007. Public health surveillance for coccidioidomycosis in Arizona. Annals of the New York Academy of Sciences, 1111: 96–102.

Swatek, F.E., 1970. Ecology of Coccidioides immitis, Mycopathotogia et Mycologia applicata, 40(1-2): 3-12.

Tamerius, J.D., and A.C. Comrie, 2011. Coccidioidomycosis incidence in Arizona predicted by seasonal precipitation, PloS ONE 6(6): e21009, doi:10.1371/journal.pone.0021009.