ECOLOGY AND GEOGRAPHY OF HUMAN MONKEYPOX CASE OCCURRENCES ACROSS AFRICA

by

CHRISTINE K. ELLIS

B. Sc., KANSAS STATE UNIVERSITY, 1988
DVM, KANSAS STATE UNIVERSITY, 1990

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF PUBLIC HEALTH

Kansas State University MPH Program (Infectious Diseases and Zoonosis)
and
Department of Diagnostic Medicine and Pathobiology
College of Veterinary Medicine

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Approved by:
Major Professor
Patricia Payne DVM PhD

Abstract

As ecological niche modeling (ENM) evolves as a tool in epidemiology and public health, refinement of occurrence data , and selection of the most appropriate and informative environmental data sets becomes increasingly important. In this report, a previous ENM analysis predicting the potential distribution of human monkeypox in Africa is reassessed using refined georeferencing criteria, and use of a more diverse set of environmental data, in order to identify environmental parameters contributing to monkeypox ecology. Significant environmental variables included annual precipitation, several temperature-related variables, net primary productivity, potential evaoptranspiration, solid moisture, soil pH, and two monthly NDVI variables. Our results emphasize the importance of selecting the most appropriate and informative environmental data for ENM analysis.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deepest gratitude to Dr. A. Townsend Peterson, University of Kansas, Natural History Museum & Biodiversity Research Center, Department of Ecology & Evolutionary Biology, Lawrence, KS for his patience and tutelage, which enabled me to develop the project, and for serving on my committee.

I would like to thank my Major Professor, Dr. Patricia Payne for the encouragement, advice and the good-natured patience and insightfulness she extended to me during the course of my study.

I would like to thank Dr. Gail Hansen and Dr. Deon van der Mewre for serving on my committee, and for their insights and encouragement.

I would like to thank Mona Papes, and Ryan Lash for their patience, guidance and encouragement in helping me understand and learn the nuances of GIS and georeferencing, which were so important to this project.

Finally, I would like to thank Gina Scott, the Instructional Services Coordinator at the Kansas State University, College of Veterinary Medicine for the help and advice she shared with me as I prepared my figures and tables for inclusion in this report.

CHAPTER 1 - Introduction

Monkeypox

Human monkeypox is a zoonotic smallpox-like febrile rash illness caused by monkeypox virus, a double-stranded DNA virus that is a member of the genus *Orthopoxvirus* (family *Poxviridae*, sub-family *Chordopoxvirinae*) along with camelpox, cowpox, ectromelia, vaccinia, and variola (smallpox)(Arita et al. 1985, Nalca et al. 2005, Breman 2000, Sale, Melski and Stratmen 2006). Poxviruses are large (130-360 kpb) viruses that replicate within the cytoplasm of infected cells via viral associated DNA-polymerase (Lewis-Jones 2004). The central part of the genome is tightly conserved (i.e., has 93.6% central homology to variola) and contains genes involved in key functions such as transcription and virus assembly (Lewis-Jones 2004). Genes located at the termini are more variable, and are associated with virus-host interactions such as host range restriction, immune system evasion, and host specificity (Lewis-Jones 2004, Stanford et al. 2007). The high level of conservation contributes to a high degree of antigenic similarity, enabling use of one virus species as a means to protect against infection by another. Examples of immunization efforts that made use of this antigenic similarity include use of cowpox virus by Edward Jenner to provide immunity against smallpox, and the use of vaccinia virus to induce protective immunity against smallpox during worldwide eradication efforts (Stanford et al. 2007).

Whole genome analysis, restriction fragment-length polymorphism analysis (RFLP), and DNA sequencing of monkeypox virus (MPXV) have identified the presence of two geographically distinct MPXV clades that are 99% identical, and demonstrate greatest diversity within the terminal regions (Parker et al. 2007, Mackett and Archard 1979, Esposito and Knight 1985, Likos et al. 2005, Reed et al. 2004). The Congo Basin clade is comprised of MPXV isolates collected in Cameroon, Republic of the Congo, Gabon and the Democratic Republic of the Congo (DRC), whereas isolates from Nigeria, Liberia, and those imported into the United States from Ghana constitute the West African clade (Esposito and Knight 1985, Likos et al. 2005, Reed et al. 2004, Sale et al. 2006). The West African clade appears to be significantly less virulent and transmissible

than the Congo Basin clade, and in general both clades appear less virulent than variola (Parker et al. 2007).

A number of clade-specific proteins and orthologues (genes that are similar to each other), found within the terminal regions, may contribute to the differences noted in disease severity and transmission via modulation of viral pathogenesis and/or host responses, and clearance of virus from infected hosts (Likos et al. 2005). For example, the West African clade contains an unique 9 amino acid epitopes that may facilitate efficient host immune recognition and clearance of West African MPXV-infected cells (Likos et al. 2005). Additionally, the West African clade does not encode a functional monkeypox inhibitor of complement enzymes (MOPICE) which may increase virus and virus-infected cell susceptibility to host-derived complement-mediated lysis, resulting in less severe disease, lower viremia, and decreased transmissibility (Parker et al. 2007, Stanford et al. 2007).

Monkeypox was initially identified as a disease of primates in 1959, when a disease outbreak occurred in a colony of cynomolus monkeys (*Macaca fasicularis*) at the State Serum Institute in Copenhagen Denmark (Arita et al. 1985, Sale et al. 2006, Parker et al. 2007, Stanford et al. 2007). Nine additional outbreaks in North America and Europe occurred over the next 10 years. In each of these outbreaks, infection resulted in a vesiculo-pustular smallpox-like disease affecting monkeys (and one South American anteater at the Rotterdam Zoo), but not the humans who handled infected animals (Arita et al. 1985, Sale et al. 2006, Parker et al. 2007, Stanford et al. 2007).

Human monkeypox was not identified as a distinct disease until 1970-1971 when successful smallpox eradication efforts in rural tropical rainforest areas of western and central Africa revealed the presence of this smallpox-like illness (Arita et al. 1985, Breman 2000, Huhn et al. 2005, Nalca et al. 2005, Parker et al. 2007, Sale et al. 2006, G. et al. 1980). During this period, 6 human cases were reported in Liberia, Sierra Leone, Nigeria, and Zaire (present-day DRC), and over the next decade, 53 additional cases would be reported throughout the countries of sub-Saharan Africa (Nalca et al. 2005). Eighty percent of these cases occurred in the DRC, with the remainder occurring in Cameroon, the Central African Republic, Gabon, Cote d'Ivoire, Liberia, Nigeria, and Sierra Leone (Likos et al. 2005, Nalca et al. 2005, Parker et al. 2007).

Concerns that monkeypox might be able to fill the niche vacated by variola lead the World Health Organization (WHO) to initiate a monkeypox surveillance program in the DRC from 1981-1986. The program identified 404 human monkeypox cases despite being hampered by the lack of a robust MPXV antibody-specific immunoassay, and the gross clinical similarity between human monkeypox and varicella zoster virus (VZV) (Arita et al. 1985, Nalca et al. 2005, Parker et al. 2007). After the surveillance program ended, reports of human monkeypox declined (13 cases reported from 1986-1992, none reported from 1993-1995) (Likos et al. 2005, Parker et al. 2007). Numbers rebounded in 1996 when 133 cases in western and central Africa were reported, followed by 511 cases in 1997. Case numbers continued to rise from 1998-2002, although lack of laboratory confirmation in some of these reports could not rule-out the possibility that some cases were caused by VZV (Nalca et al. 2005, Parker et al. 2007).

In the spring of 2003, monkeypox emerged in the Western Hemisphere when a cluster of human cases were reported in the United States (Huhn et al. 2005, Nalca et al. 2005, Sbrana et al. 2007, Stanford et al. 2007). By the end of the outbreak, 72 cases (37 confirmed) were reported in 6 Midwestern States (Illinois, Indiana, Kansas, Missouri, Ohio, and Wisconsin) (Huhn et al. 2005). While source of the virus was determined to be a consignment of West African rodents imported from Ghana, the source of human exposure was determined to be captive prairie dogs (*Cynomys* spp.) that had been housed in close proximity to the infected West African rodents prior to entering the pet market (Nalca et al. 2005, Parker et al. 2007, Sbrana et al. 2007, Reynolds et al. 2006). In this outbreak, transmission of the virus to humans occurred primarily via direct contact with infected prairie dogs, and by indirect exposure via aerosol or fomites (Parker et al. 2007, Reynolds et al. 2006). There were no instances of human-to-human transmission. (Reynolds et al. 2006)

The clinical course of disease observed during this outbreak was less severe than that described in Central Africa, in that most of the U. S. cases exhibited a mild self-limiting flu-like illness with few vesicular lesions, and fewer children under the age of 10 years of age were infected (Likos et al. 2005, Nalca et al. 2005, Reynolds et al. 2006, Sbrana et al. 2007). The decreased severity of disease was likely due to several factors including: a higher natural resistance in the U. S. population, a healthier population of

people in the U.S., better access to quality health care, and variability in the pathogenicity of the virus involved in the outbreak (Chen et al. 2005, Likos et al. 2005). Variability in pathogenicity was eventually confirmed when RFLP and DNA sequencing analysis identified the viral strain as a member of the less virulent West African clade (Chen et al. 2005, Likos et al. 2005).

The clinical presentation of human monkeypox is similar to that of smallpox (Arita et al. 1985, Nalca et al. 2005, Reynolds et al. 2006, Sbrana et al. 2007). Onset of clinical signs begins after a 10-14 day incubation period followed by a 1-3 day prodromal illness characterized by fever, malaise, lymphadenopathy, and upper respiratory tract illness (Nalca et al. 2005, Sbrana et al. 2007). Over a 2-4 week period, a rash comprised of 0.5-1.0 cm diameter lesions appears and progresses through macular, papular, vesicular, and pustular stages followed by umblication, scabbing, and desquamation (Arita et al. 1985, Parker et al. 2007, Sbrana et al. 2007). The rash typically appears on the trunk, and spreads centrifugally to the limbs, hands and feet, and occasionally to the oral mucosa and genitalia (Parker et al. 2007, Nalca et al. 2005). Unilateral or bilateral lymphadenopathy involving submandibular, inguinal, or axillary lymph nodes is observed in 90% of patients and may be considered a key feature differentiating human monkeypox from variola and VZV (Arita et al. 1985, Nalca et al. 2005, Parker et al. 2007). Complications may include recurrent fever, secondary skin or soft tissue infection, coagulation disorders, pneumonitis, ocular lesions, encephalitis, and multi-organ failure which are often indicative of a fatal outcome (mortality 10%) (Nalca et al. 2005, Parker et al. 2007, Sbrana et al. 2007).

The total epidemiologic range of human monkeypox has not been definitively identified, however, it is likely to parallel the ecological niche occupied by the zoonotic reservoir host(s), which are unknown at this time (Breman 2000, Khodakevich, Jezek and Kinzanzka 1986, Khodakevich, Jezek and Messinger 1988, Learned et al. 2005, Nalca et al. 2005, Sale et al. 2006, Stanford et al. 2007). Most data indicates that primates function only as incidental hosts, and that reservoir hosts are most likely to be one or more rodent species indigenous to west and central Africa (Stanford et al. 2007, Breman 2000, Khodakevich et al. 1986, Khodakevich et al. 1988, Nalca et al. 2005, Sale et al. 2006).

Ecological investigations conducted in the lowland tropical forests of Central and West Africa by the Centers for Disease Control and Prevention (CDC), and the WHO during the 1970s-1980s identified several species of animals capable of mounting an immune response to presumed MPXV infection (Table 1) (Parker et al. 2007). Seropositive animals were not confined to one ecological strata; 40% were arboreal, 40% were semi-terrestrial, and 20% were terrestrial, suggesting that the sylvatic cycle of monkeypox may be comprised of interactions between incidental and reservoir hosts occupying overlapping ecological niches within the lowland tropical rain forests (Parker et al. 2007). Monkeypox virus has only been isolated from African animal species on two occasions. In 1985, MPXV was isolated from a rope squirrel (*Funisciurus anerythrus*) captured in the DRC, and in 2003, in association with the U. S. outbreak, virus was isolated from rope squirrels, a Gambian giant pouched rat (*Cricetomys gambianus*), and African dormice **(***Graphiurus* spp.) (Table 1) (Breman 2000, Khodakevich et al. 1986, Parker et al. 2007, G. et al. 1980).

GENUS SPECIES	**COMMON NAME**	**TESTING METHODOLOGY**	**HABITAT**
Rodents			
Cricetomys emini	African/Emin giant pouched rat	Serology	Terrestrial
Cricetomys gambianus	Gambian giant pouched rat	Serology PCR	Terrestrial
Funisciurus anerythrus	Thomas' rope squirrel	Serology Virus isolation	Semi-terrestrial
Funisciurus congicus	Congo rope squirrel	Serology	Semi-terrestrial
Funisciurus ilsabella	Lady Burton's rope squirrel	Serology	Semi-terrestrial
Funisciurus lemniscatus	Ribboned rope squirrel	Serology	Semi-terrestrial
Graphiurus spp.	African dormice	Serology PCR	Terrestrial
Heliosciurus gambianus	Gambian sun squirrel	Serology	Arboreal
Heliosciurus rufobrachium	Red-legged sun squirrel	Serology	Terrestrial
Lophuromys sikapusi	Rusty-bellied rat	Serology	Terrestrial
Mastomys gr. Coucha	Multimammate rat	Serology	Terrestrial
Petrodromus tetradactylus	Four-toed elephant shrew/ Sengi	Serology	Terrestrial
Protoxerus strangeri	Forest giant squirrel	Serology	Semi-terrestrial
Miscellaneous Mammals			
Atherurus africanus	African brush-tailed porcupine	Serology	Semi-terrestrial
Cephalophus monticola	Blue duiker	Serology	Terrestrial
Sus scrofa	Pig	Serology	Terrestrial
Non-human primates			
Allenopithecus nigrovirdis	Allen's swamp monkey	Serology	Semi-terrestrial
Cercocebus galeritus	Agile / Crested mangabey	Serology	Semi-terrestrial
Cercopithecus aethiops	Grivet	Serology	Semi-terrestrial

Cercopithecus ascanius	Red-tailed monkey	Serology	Arboreal
Cercopithecus diana	Diana monkey	Serology	Arboreal
Cercopithecus mona	Mona monkey	Serology	Semi-terrestrial
Cercopithecus nictitans	Greater white-nosed monkey	Serology	Arboreal
Cercopithecus petaurista	Lesser white-nosed monkey	Serology	Arboreal
Colobus badius	Red colobus monkey	Serology	Arboreal
Pan troglodytes	Chimpanzee	Not Recorded	Semi-terrestrial
Avian			
	Calao (toucan) *	Serology	
	Touraco *	Serology	
	* denoted by colloquial term only		

Table 1: Animals identified as capable of mounting an immune response to presumed monkeypox infection (Parker et al. 2007, Arita et al. 1985, Breman et al. 1980).

The capability of other animal or insect species to function as reservoir or vector hosts of monkeypox has not been fully investigated. For example, there have been few studies evaluating the potential role insects might play in the sylvatic cycle of monkeypox (Parker et al. 2007). The potential reservoir host capability of animal species indigenous to locales outside the presumed range of monkeypox must be considered as well; PCR testing of animal tissues during the 2003 U. S. outbreak demonstrated that rodents such as hamsters (*Circetus* spp.), gerbils (*Gerbillus* spp.), and chinchillas (*Chinchilla* spp.) were capable of serving as potential hosts (Parker et al. 2007).

The epidemiology of human monkeypox continues to evolve, with much remaining unknown about its geographic distribution and ecology. The virus appears to be endemic to tropical rainforested regions of West and Central Africa, with most cases occurring in the Congo Basin (Arita et al. 1985, Huhn et al. 2005, Learned et al. 2005, Reynolds et al. 2006, Sbrana et al. 2007). Notable differences in the epidemiological and clinical features of human monkeypox are present when comparing disease caused by the Congo Basin *versus* West African clades. In Central Africa higher case numbers, and increased morbidity, mortality, and human-to-human transmission are reported, whereas disease in West African appears attenuated and less transmissible (Arita et al. 1985, G. et al. 1980, Chen et al. 2005, Foster et al. 1972, Hutin et al. 2001, Janseghers et al. 1984, Ladnyj, Ziegler and Kima 1972, Levine et al. 2007, Likos et al. 2005, Meyer et al. 1991).

Human monkeypox shares some clinical features with variola, vaccinia, and cowpox viruses, but differs epidemiologically in its transmission and case fatality characteristics. Human disease occurs sporadically in clusters, and is thought to occur

primarily via direct contact with infected animals or infected animal tissues (~ 72% of cases) in both West and Central Africa (Breman 2000, Jezek, Gromyko and Szczeniowski 1983, Jezek et al. 1986, Jezek et al. 1988, Khodakevich et al. 1988, Parker et al. 2007, Sale et al. 2006). The incidence of human-to-human transmission (the secondary attack rate) is low (~ 9%), is not sustainable, and is thought to occur during the febrile prodrome via direct contact, respiratory aerosol, or contact with body fluids (Breman 2000, Jezek et al. 1983, Jezek et al. 1986, Jezek et al. 1988, Khodakevich et al. 1988, Nalca et al. 2005, Reynolds et al. 2006, Sale et al. 2006, Stanford et al. 2007). There is some evidence that the secondary attack rate in the DRC has increased over the past 30 years, however, it remains far below that of variola (~ 60%), and mathematical models based on human monkeypox occurrence in the DRC imply that MPXV is not capable of indefinite transfer among unvaccinated humans without zoonotic amplification (Chen et al. 2005). This finding is important, because it indicates that MPXV is not likely to fill the niche formerly occupied by variola (Likos et al. 2005, Parker et al. 2007).

The case fatality rate of human monkeypox is lower than that of smallpox (~1.5-17% *versus* 17-30%, respectively), with greatest mortality reported in Central Africa (Chen et al. 2005, Huhn et al. 2005). Most cases of infection and the highest mortality rate occur in children less than 10 years old, which could be attributed in part to the emergence of a population of individuals born after eradication of smallpox (Nalca et al. 2005, Stanford et al. 2007). Smallpox vaccination confers ~ 85% cross-immunity against MPXV for 3-19 years post-immunization; it has been hypothesized that the decreased incidence of human monkeypox in west and central Africa in the years immediately following the global smallpox eradication campaign may have been related to the widespread administration of smallpox vaccine during that period (Learned et al. 2005, Nalca et al. 2005, Parker et al. 2007).

The increasing incidences of human monkeypox, its clinical similarities to variola, and inclusion of MPXV on the select list of biological agents considered possible agents of bioterrorism, make this virus the most significant orthopoxvirus infection of man with regards to surveillance and research (Breman 2000, Chen et al. 2005, Agriculture 2002). A combination of factors are likely responsible for the increase in incidence and may include: waning immunity among individual who were vaccinated

during the smallpox eradication campaign, increased numbers of susceptible children born after discontinuation of smallpox vaccination, increased dependence on hunting for food (bush meat) in monkeypox endemic localities, encroachment of humans into the ecological niches of reservoir host(s), ecosystem degradation, and increased susceptibility of humans due to poverty, socioeconomic variables, inadequate living conditions, substandard nutrition, and the presence of co-infections and/or parasitism (Arita et al. 1985, Chen et al. 2005, Nalca et al. 2005, Parker et al. 2007).

Because human monkeypox is a zoonotic disease it cannot be considered an eradicable disease (Arita et al. 1985, Reynolds et al. 2002). The broad host range suggests that this virus is capable of adapting to new hosts within its endemic range, and within new regions as well, with the 2003 U. S. outbreak serving as an eloquent example of this capability (Parker et al. 2007). Future efforts to develop better understanding of the ecology of monkeypox would ideally focus on documentation of its ecological niche, and the environmental parameters associated with occurrence in western and central Africa historically and in the event of climate change. Such information may prove valuable in identification of reservoir and incidental host(s) and naïve environments potentially capable of sustaining monkeypox should introduction occur. Additional surveillance would be beneficial in determining whether or not occurrence of human monkeypox constitutes an outbreak *versus* an endemic disease occurrence, and whether incidence of disease is proportional to waning immunity to smallpox vaccination, and environmental and socioeconomic factors.

Ecological Niche Modeling

The ecological niche of a species is defined as the set of environmental conditions capable of maintaining populations without immigration from other areas (Grinnell 1917, Grinnell 1924, Levine et al. 2007, Peterson et al. 2005, Townsend, Sanchez-Cordera and Martinez-Meyer 2005, Peterson and Nakazawa 2008). An ecological niche model (ENM) can be described as the probability distribution of a species as defined by a set of environmental variables (i.e., annual temperature, annual precipitation, land cover, etc.) and species localities of known occurrence (Peterson 2001, Santiago 2005). This probability distribution can then be integrated into a geographic information system

(GIS) to identify geographic regions containing environmental conditions that could support the species (Santiago 2005, Araujo and Guisan 2006, Elith et al. 2006, Peterson 2007, Sweeney et al. 2006, Papes and Gaubert 2007).

Application of ENMs have proven useful within the fields of biogeographical research, conservation biology, ecology, paleoecology, wildlife conservation and management, and recently in the field of spatial epidemiology, by providing predictive pictures of past, current, and potential species distributions within geographic locales deemed ecologically fit for the species, even if the species is not present (Araujo and Guisan 2006, Costa, Peterson and Beard 2002, Peterson 2003, Peterson 2006, Santiago 2005, Peterson and Nakazawa 2008, Peterson, Carroll and N. 2004) Additional applications of ENMs include examination of the influence of geographic alterations and environmental shifts (i.e., climate change, or human-derived land use) on species distributions, evaluation of potential geographic outcomes of species interactions, and identification of geographic areas susceptible to species invasion (Costa et al. 2002, Santiago 2005, Peterson 2006, Peterson 2007, Ward 2007).

Within the field of spatial epidemiology, ENMs have the potential to be useful in identification of ecologic, geographic and spatial characteristics of disease occurrence and transmission (Araujo and Guisan 2006, Costa et al. 2002, Peterson et al. 2004, Peterson 2006, Peterson 2007, Levine et al. 2007, Peterson and Nakazawa 2008, Peterson et al. 2005). Traditional methodologies used to evaluate the geographic risk of disease transmission often focus broadly on the overall distribution of cases within a geographic space, with little emphasis placed on the transmission system involved (Peterson 2007). However, disease occurrence is often comprised of complex interactions between multiple species (host, pathogen, vector), each distributed according to its own ecological potential across the landscape, therefore, it is important to consider that the geographical distribution of disease is an epidemiological event comprised of the interactions of each participating species' ecology (Peterson 2006). In some situations, the ecological factors related to disease transmission and occurrence may not be fully understandable until the individual ecologies of the vector, host, and pathogen are characterized independently and comparatively (Peterson 2006, Peterson 2007).

Application of ENMs to disease may include investigation of landscapes for areas that meet the ecological requirements of the species involved in disease occurrence, identification of the components involved in transmission cycles, and construction of models of predicting species invasion (Peterson 2006, Santiago 2005, Costa et al. 2002, Peterson 2003, Peterson 2007, Sweeney et al. 2006). ENMs may be used to investigate landscapes that meet the ecologic requirements of the species involved in disease occurrence by identifying both the true geographic range and potential areas of occurrence that are suitable but currently naïve (Peterson 2006, Santiago 2005). Application of such models allows investigation into the ecology of the transmission chain, identification of landscapes suitable for persistence after introduction or invasion, development of interventions that could preclude introduction of a pathogen into a naïve landscape, and implementation of programs designed to moderate the impact of invasion or establishment (Peterson 2006, Santiago 2005, Costa et al. 2002, Peterson 2003, Peterson 2007). ENMs may also be used to model the components within a disease transmission cycle (i.e., case distribution, and pathogen, reservoir host, or vector distributions) in order to assemble a comprehensive ENM representing the broad geographic picture of the transmission system, and as a means to identify unknown elements within that system (Costa et al. 2002, Peterson 2003, Peterson et al. 2004, Peterson 2006, Peterson et al. 2006a, Peterson 2007). For example, the basic ecological requirements (and in some instances, the identity) of all of the species participating in a disease transmission cycle are often unknown, and by modeling the "known" (geographic case occurrence, the ecological requirements and geographic potential of known species), inference of unknown, likely, or suspect species may be acquired (Costa et al. 2002, Peterson 2003, Peterson et al. 2004, Peterson 2006, Peterson et al. 2006a, Peterson 2007, Santiago 2005, Sweeney et al. 2006).

Strengths of ENM include the ability to characterize the ecological and geographic requirements of species in "real-world" space, and independence from any specific landscape. This independence gives ENMs the versatility to objectively identify areas of potential distribution in any landscape (sampled or unsampled, known, potential or changing). (Peterson 2006) In addition, improvements to existing spatial epidemiological methodologies used to research geographic patterns of disease

transmission and disease risk can be made through the use of ENMs (Peterson 2007). One valuable contribution is the ability of ENMs to achieve fine scale resolution of distributions, as compared to commonly used spatial epidemiology techniques (Peterson 2006, Peterson 2007).

Issues that remain to be addressed in the application of ENMs to disease systems include biased reporting of occurrence, small sample sizes, lack of detailed geographic or ecologic analysis, and inappropriate matching of temporal and spatial scales (Peterson 2006, Peterson 2007). Other uncertainties that could impact the use of ENMs include reliance on presence-only data which may lead to over-fitting biases, and uneven sampling effort (small sample size, poor sampling strategy, erroneous locality descriptions) resulting in ENMs affected by omission (under-prediction) error (known areas of presence that are predicted as absent) or commission (over-prediction) error (areas of absence that are predicted as present) (Araujo and Guisan 2006, Papes and Gaubert 2007, Stockwell and Peterson 2003). ENMs based on low numbers of presence-only data (10-20 points) have been shown to be reliably accurate, however, the use of such data is not ideal (but is sometime unavoidable) and can exacerbate the biases inherent in the use of presence only data, resulting in an ENM that is not fully representative of the ecological niche (Stockwell and Peterson 2003, Peterson 2007, Papes and Gaubert 2007)..

The quality, source, and quantity of environmental datasets (i.e., climate data, remotely sensed environmental data) used to construct an ENM may affect the outcome as well, because inclusion of too little environmental data may result in under-prediction, while too much may result in over-prediction. (Stockwell and Peterson 2003). Climate data is important in model construction because it provides long temporal applicability, while remotely sensed environmental data (i.e., land surface reflectance, landform, substrate, topography, vegetative indices) contributes to the model by measuring different aspects within the ecological landscape at fine spatial resolutions that can be considered representative of real species presence or absence (Peterson 2006, Peterson 2007, Peterson and Nakazawa 2008, Townsend et al. 2005, Costa et al. 2002). After identification of the most appropriate environmental datasets, selection of appropriate data variables from within each environmental dataset is equally important in order to

develop ENMs that are "best fit" and capable of extrapolation across space, time and onto novel landscapes (Peterson and Nakazawa 2008, Peterson 2006, Peterson 2007). Test ENMs characterizing the relationship between the occurrence data, and grouped and individual environmental variables should be evaluated at the beginning of the modeling process in order to identify those environmental variables most significant for production of the final model. Final characterization of the relationship between the occurrence data and the most appropriate environmental variables can be accomplished via a variety of methods including: range-based rules, additive and linear statistical models (i.e., linear regression), distance- and factor-based methods, and machine learning computing approaches.

Machine learning computing methodologies have proven to be robust and reliable at defining the complex relationships between occurrence data and environmental variables with maximum flexibility and less bias (Peterson 2007). Maxent (version 3.0, www.cs.princeton.edu/~schapire/maxent/) is an example of a general purpose, maximum entropy-based machine learning method used to estimate the probability distribution for species occurrence (Phillips, Dudik and Schapire 2004) Maxent constructs ENMs based on the environmental characteristics of presence-only occurrence data and 10,000 random background points representing areas of non-occurrence (pseudoabsence) within the study area (Elith et al. 2006). Output is in the form of ASCII raster grids, which may then be imported into GIS programs for analysis.

Maxent estimates the probability distribution of a species within the study landscape by identifying the distribution of maximum entropy (i.e., the probability distribution closest to uniform) subject to the constraint that the expected value of each environmental variable within the estimated probability distribution should match its empirical average (Phillips, Anderson and Schapire 2006). Predicted values within the estimated distribution are initially represented as raw probabilities that sum to unity; consequently these values tend to be low when the extent of the analysis is large (Elith et al. 2006, Phillips et al. 2004). Maxent results may also be presented as cumulative values, whereby each cell in the output raster is given a value equal to its assigned probability plus the sum of all lower probabilities. Thus, a value of 100 indicates highest suitability, whereas values close to or equal to 0 are considered unsuitable (Phillips et al.

2004, Hernandez et al. 2006, Peterson et al. 2006b, Peterson, Papes and Eaton 2007). To avoid overfitting, Maxent utilizes a smoothing feature called regularization (a relaxation function) to constrain estimated distributions, so that the average value for a given predictor remains close to the empirical average, and within the empirical error limits (Phillips et al. 2004, Hernandez et al. 2006).

Advantages of Maxent over other machine learning tools include the speed and simplicity of the software implementation, and its ability to: make predictions when incomplete data is available; estimate probability distributions that are spread out given the constraints derived from the available data; use both categorical and continuous environmental data; and produce detailed predictions based on the continuous nature of the resulting models (Papes and Gaubert 2007).

CHAPTER 2 - Materials and Methods

Human Monkeypox Occurrence Data

Locations of known case occurrences of human monkeypox in endemic regions in West and Central Africa were compiled from outbreak investigation and surveillance data provided by the CDC and the WHO (Levine et al. 2007). For this study, a human monkeypox case was defined as a published reported case or a non-redundant unpublished case confirmed by laboratory evidence of disease. Laboratory detection methods used to classify human monkeypox cases recorded between 1970-1986 by the WHO included electron microscopy (EM), virus culture, and serology, whereas polymerase chain reaction (PCR), EM, and tissue culture were used for case definition by CDC (Learned et al. 2005, Levine et al. 2007).

Overall, 404 human monkeypox cases documented geographically with variable degrees of specificity (i.e., to country, region, district/zone, municipality, or specific locality) were available for inclusion in this study. Geographic coordinates were assigned to cases based on municipality and specific locality, based on consultation of the Alexandria Digital Library Gazetter (www.alexandria.ucsb.edu), National Geospatial Intelligence Agency Geographic Names Databases (www.gnswww.nga.mil/geonames/GNS/index), and electronic data published with the Rand McNally New Millennium World Atlas (Rand McNally, 1988). The MaNIS point-radius method of georeferencing was used to assess spatial uncertainty in the geographic referencing of each occurrence point to account error associated with the spatial extent of the named place, uncertainty of directions, and uncertainty of distances (Wieczorek, Guo and Hijmans 2004, Peterson 2008b). We restricted our analyses to sites that could be georeferenced with a spatial precision finer than 10 km^2. Because most of the cases were poorly described geographically, only 216 occurrence localities could be used. Redundant case occurrences (i.e., cases with identical coordinates) were removed, leaving 139 occurrences available for analysis.

Environmental Data Sets

Environmental data were drawn from 4 principal sources. Climatic data were drawn from the WorldClim archive (www.worldclim.org), a climate database containing global climate data interpolated from weather station data for 1950-2000 at 10' (~344 km^2) spatial resolution (Hijmans et al., 2005). Nineteen 'bioclimatic' variables were initially explored: annual mean temperature; annual precipitation; isothermality; maximum temperature of the warmest month; minimum temperature of the coldest month; mean diurnal temperature range; mean temperature and precipitation of the coldest, driest, warmest, and wettest quarters; precipitation of the driest and wettest months; precipitation seasonality, temperature annual range, and temperature seasonality (Hijmans et al., 2005). Data sets summarizing soil and vegetation characteristics were obtained from the GeoData Portal (United Nations Environment Programme; www.geodata.grid.unep.ch/data), including data layers summarizing net primary productivity (NPP), potential evapotranspiration (pevap), soil carbon, soil moisture, and soil pH. Topographic data were obtained from the U.S. Geological Survey's Hydro-1K (www.edc.usgs.gov/products/elevation/gtopo30/hydro/indexdigital) digital elevation model, including aspect, compound topographic index, flow accumulation, and slope. Finally, we used composite Normalized Difference Vegetation Index (NDVI) coverages derived from the Advanced Very High Resolution Radiometer (AVHRR) satellite (University of Maryland, www.glcf.umiacs.umd.edu/index.html) to summarize monthly photosynthetic mass during April 1992 – March 1993 as an exemplar year. All datasets were resampled to 10 km spatial resolution for analysis to match the appropriate spatial precision of the case occurrences.

Ecological Niche Modeling

The ecological niche of a species can be defined as the set of environmental conditions under which it is able to maintain populations without immigrational subsidy (Peterson et al. 2006b, Grinnell 1917, Grinnell 1924). Ecological niches can be estimated by integrating information on spatial occurrences of the species with relevant raster data layers summarizing aspects of the environment (Araujo and Guisan 2006). Once developed, niche models can be used to identify suitable areas for populations of the

species, effectively creating potential distribution maps for the species (Austin, Nicholls and Marguleis 1990, Costa et al. 2002, Peterson 2001, Peterson, Stockwell and Kluza 2002b, Peterson, Ball and Cohoon 2002a, Peterson 2003, Peterson 2006, Peterson et al. 2006b).

We used Maxent to generate the ecological niche models in this study. Maxent is a general-purpose, maximum entropy-based evolutionary-computing tool for inferring niche dimensions (Phillips et al. 2004). Maxent is used to estimate the probability distribution for species' occurrences by identifying the distribution of maximum entropy (i.e., a probability distribution closest to uniform), subject to the constraint that the expected value of each environmental variable within the estimated distribution should match its empirical average (Phillips et al. 2006). Maxent builds niche models based on environmental characteristics of presence-only occurrence data and 10,000 randomly chosen background points representing areas of non-occurrence (pseudoabsence) across the study area (Elith et al. 2006).

Predictions generated for each grid cell by Maxent are initially raw probabilities that sum to unity, and consequently are low when the extent of analysis is large. Maxent results are more commonly presented as cumulative values (i.e., each cell receives a value equal to its assigned probability plus the sum of all lower probabilities), wherein a value of 100 indicates highest suitability and values close to 0 would be unsuitable (Phillips et al. 2004, Hernandez et al. 2006, Peterson et al. 2007). To avoid overfitting, Maxent employs a smoothing feature called regularization (a relaxation function) to constrain estimated distributions, such that the average value for a given predictor remains within the empirical error boundaries and close to the empirical average (Phillips et al. 2004, Hernandez et al. 2006). Maxent output is in the form of ASCII raster grids, which are then imported into GIS programs for analysis.

Identification of Key Environmental Factors

To assess importance of ecological variables, we used two Maxent model performance analyses based on jackknife analysis, in which we omitted layers from our model systematically to assess their importance in determining model quality. The first analysis grouped the ecological parameters into 10 sets, including land cover,

precipitation, productivity (NPP), seasonal NDVI data (i.e., Fall, Spring, Summer, Winter), soils, temperature, and topography to obtain a broad overview of the importance of general classes of ecological variables. After initial testing and reduction on *groups* of variables, a second analysis assessed the importance of *individual* variables within the remaining groups. Initially 19 bioclimatic variables, 12 NDVI (monthly) variables, 5 soil and vegetation variables, and 5 topographic variables were used.

As such, in each analysis, for N layers (N representing the total number of layers), we developed N niche models, then systematically omitted one layer at a time from each model to assess its significance (i.e., generated N-1 models) (Peterson and Cohoon, 1999). We measured Maxent model performance as test gain (random 50% testing) for all analyses omitting each suite of variables or individual variable, and for each suite of variables and each individual variable alone (Peterson and Cohoon 1999, Phillips et al. 2006). Variables were ranked in order of significance based on these analyses and variables appearing important were selected for construction of the final model.

Characterization of Ecological Niches

Finally, we explored the distribution of the monkeypox ecological niche in environmental space. We developed a final Maxent ecological niche model based on all available occurrence data and all environmental dimensions that had proven informative in the jackknife tests. This model prediction was then combined (Grid Combine option, ArcGIS, version 9.2) with the environmental coverages on which it was based to create a raster GIS coverage with an associated attributes table summarizing the predictions and all combinations of environmental conditions. This table was then exported for visualization as a map image, and for analysis of variables using bivariate plots.

CHAPTER 3 - Results

The human monkeypox occurrence data set used to develop the ecological niche models consisted of 139 localities, including 2 from West Africa and 137 from Central Africa (Figure 1). An ecological niche model based on this occurrence information predicted a potential distribution extending across most of the humid tropical evergreen forest areas of Africa. Favorable habitat was identified in 18 African countries (Angola, Benin, Burundi, Cameroon, Central African Republic, Côte d'Ivoire, DRC, Equatorial Guinea, Gabon, Ghana, Liberia, Nigeria, Republic of the Congo, Rwanda, Sierra Leone, Togo, Tanzania, Uganda. A geographic break in the relatively continuous distribution of the disease was noted between western Cameroon and eastern Nigeria (Figure 1, see arrow).

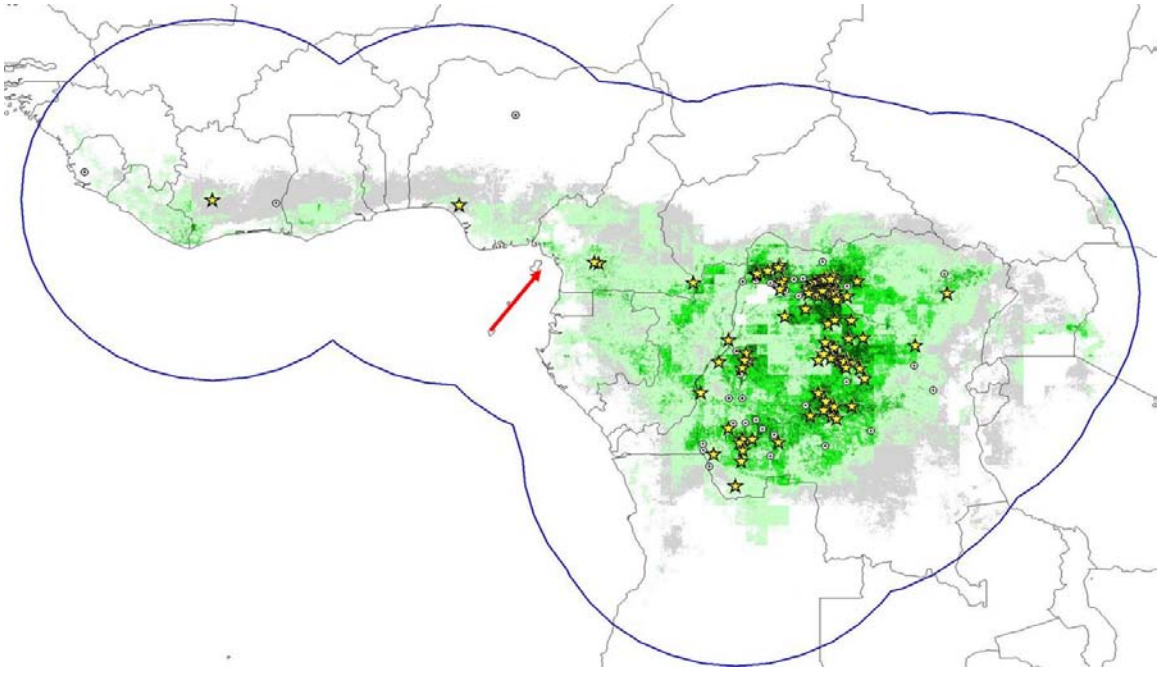


Figure 1: Ecological niche model enviornmental space predicted favorable for human monkeypox occurrence in West and Central Africa. Stars depict 139 known occurrences, dotted circles represent 216 original georeferenced localities. Shaded areas represent areas of predicted favorable environment with gray shaded areas

representing areas of lowest suitablity and dark green areas representing areas of greatest suitablity confined by a 1000 km buffer within which the analyses were developed. Red arrow indicates the geographic break in monkeypox distribution in eastern Niberia and western Cameroon.

In the first jackknife analysis, the explanatory power of each suite of variables was of greatest interest. Overall, we found that temperature variables consistently had the best explanatory power (i.e., they produced the best predictions when used alone and had the most negative effects when omitted from analysis); and that precipitation, productivity, and soil information also had some explanatory power (Figure 2). In contrast, land cover, seasonal NDVI, and topographic variables had less explanatory power (Figure 2); overall, topographic variables as a group had the lowest predictive power when analyzed in isolation.

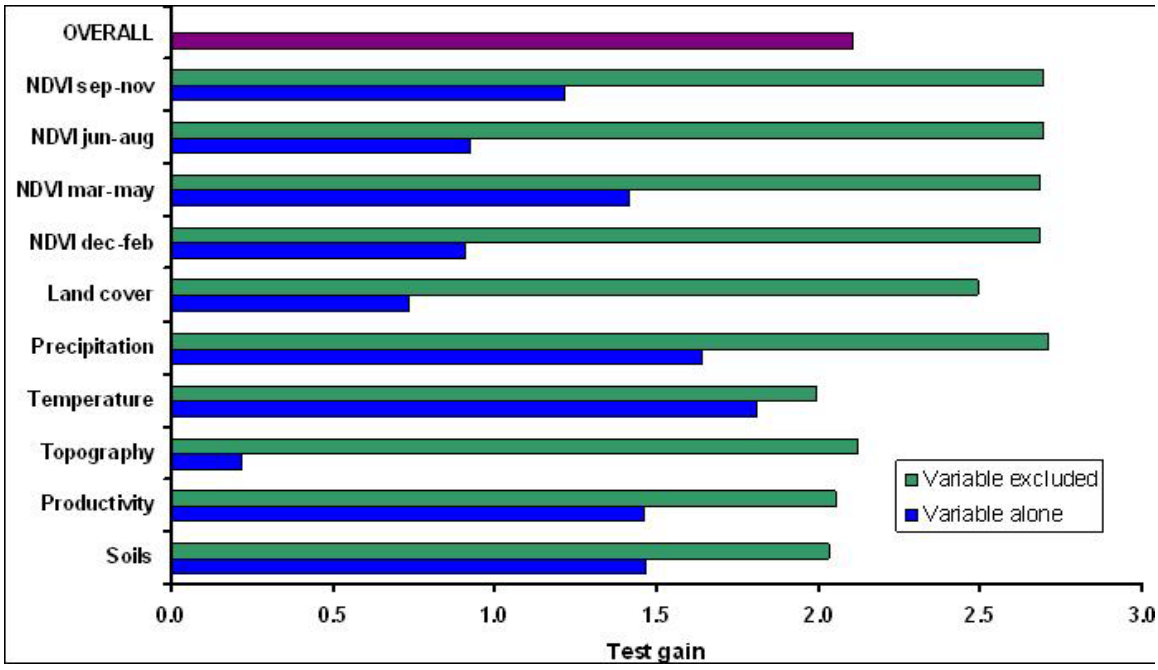


Figure 2: Summary of influences of major suites of environmental variables on predictivity of monkeypox occurrences across Africa based on Maxent models of the monkeypox ecological niche.

Removing topographic and landcover variables from consideration, we used the second jackknife manipulation to analyze the contributions of the remaining 24 individual variables to model predictivity (Figure 3). Here, net primary productivity had the greatest explanatory ability, followed by the climatic variables, soil variables, and

potential evapotranspiration. The monthly NDVI variables contained the lowest explanatory ability. All 24 variables were included in final model construction based on comparison of the overall test gain and the test gain of each N-1 data set i.e., exclusion of any variable resulted in a test gain lower that that of the overall model).

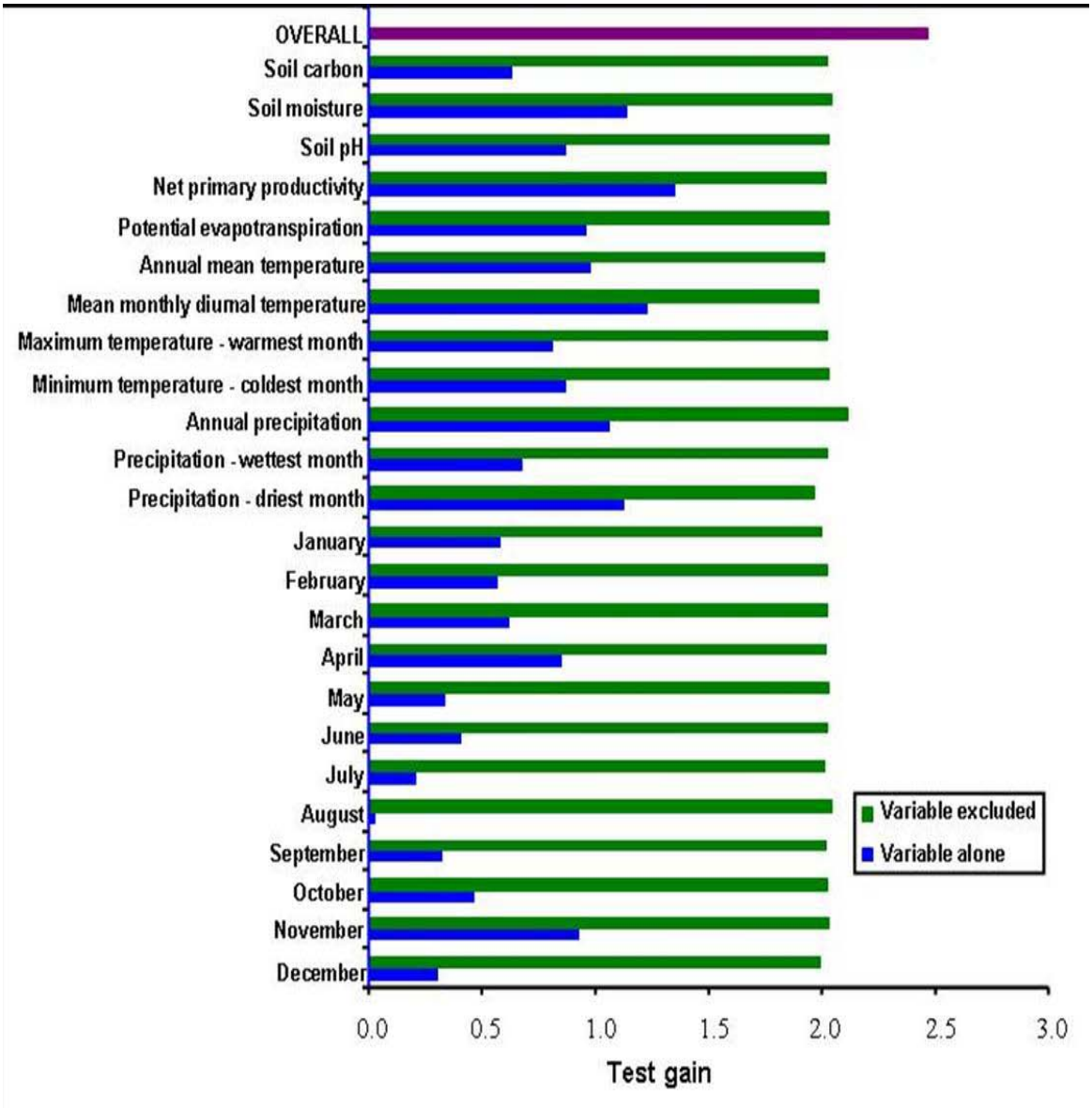
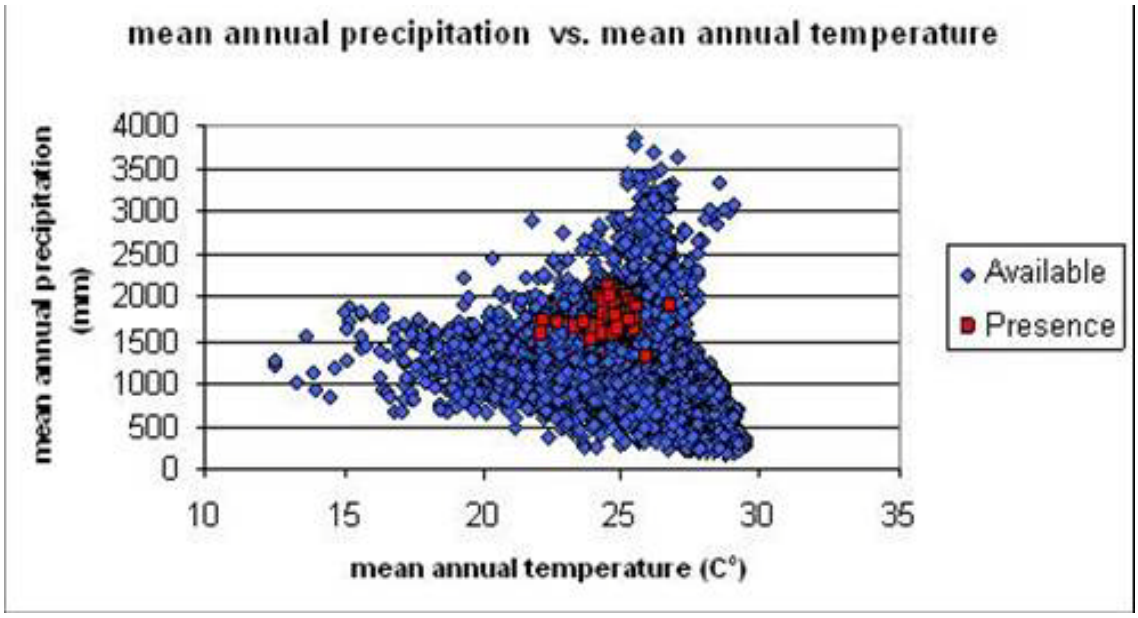


Figure 3: Summary of influences of individual environmental variables on predictivity of monkeypox occurrences across Africa based on Maxent models of the monkeypox ecological niche. NDVI variables are designated by month.

Bivariate plots were constructed using all environmental variables used in model construction to visualize areas and conditions modeled as suitable and unsuitable for monkeypox transmission. For example, monkeypox occurrence is most likely to occur in areas where mean annual precipitation ranges from 1500-2200mm , mean annual temperature ranges from approximately 21-26 ^{o}C (Figure 4). All environmental variables were examined using this methodology (Table 2)



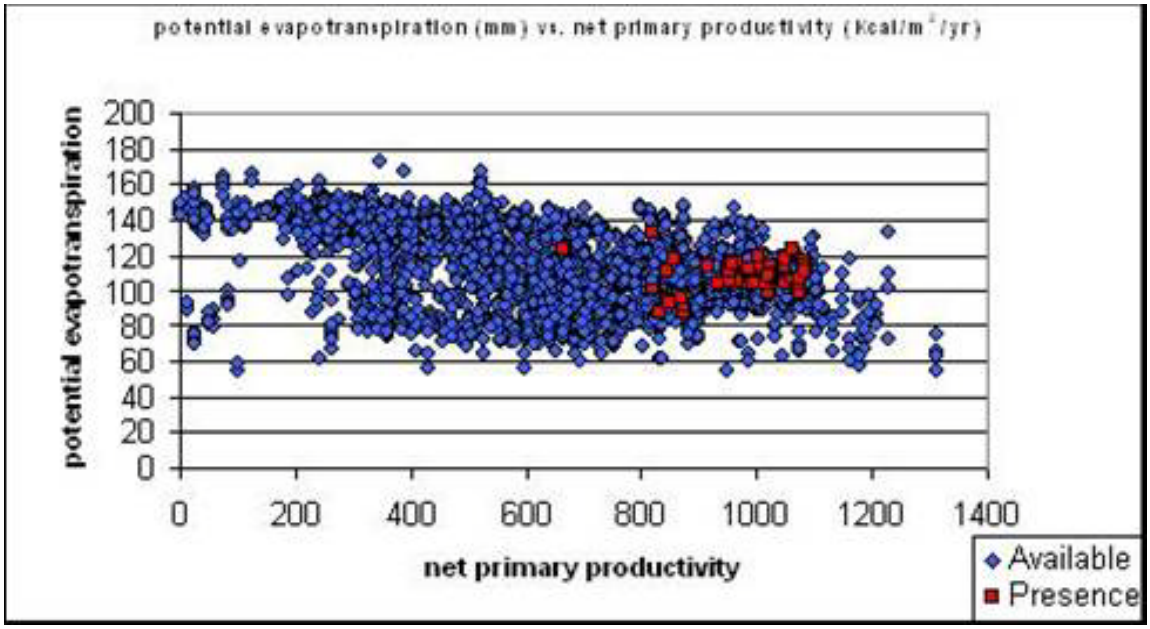


Figure 4: Bivariate plots demonstrating the distributions of monkeypox within environmental dimensions. Blue squares represent total environmental availability across West and Central Africa. Red squares represent environmental areas modeled as appropriate for monkeypox.

	Areas of Known Occurrence	**Areas of Predicted Favorable Environment**
Mean annual precipitation	1500-210 mm	500-3500 mm
Precipitation of wettest month	2000-2500 mm	600-4500 mm
Mean annual temperature	21-26 ^{o}C	15-30 ^{o}C
Maximum temperature warmest month	30-34 ^{o}C	25-42 ^{o}C
Potential evapotranspiration	80-120 mm	60-160 mm
Net primary productivity	800-1100 Kcal/m^2/yr	100-1200 Kcal/m^2/yr
Soil pH	5-6	5-7.5
Soil carbon	4-10 %	1-11 %

Table 2: Significant environmental variables as determined by bivariage plot analysis

These visualizations of conditions of predicted presence and absence can also be restricted to specific zones to characterize barriers of dispersal potentially limiting a species' distribution in a particular area. For instance, across three transects crossing different portions of the monkeypox range boundary (Figure 5), the limitation of monkeypox to areas of moderate temperature and high precipitation is consistent, but the distribution of the disease relative to November NDVI and soil moisture is variable or not separable from one transect to the next (Figure 6).

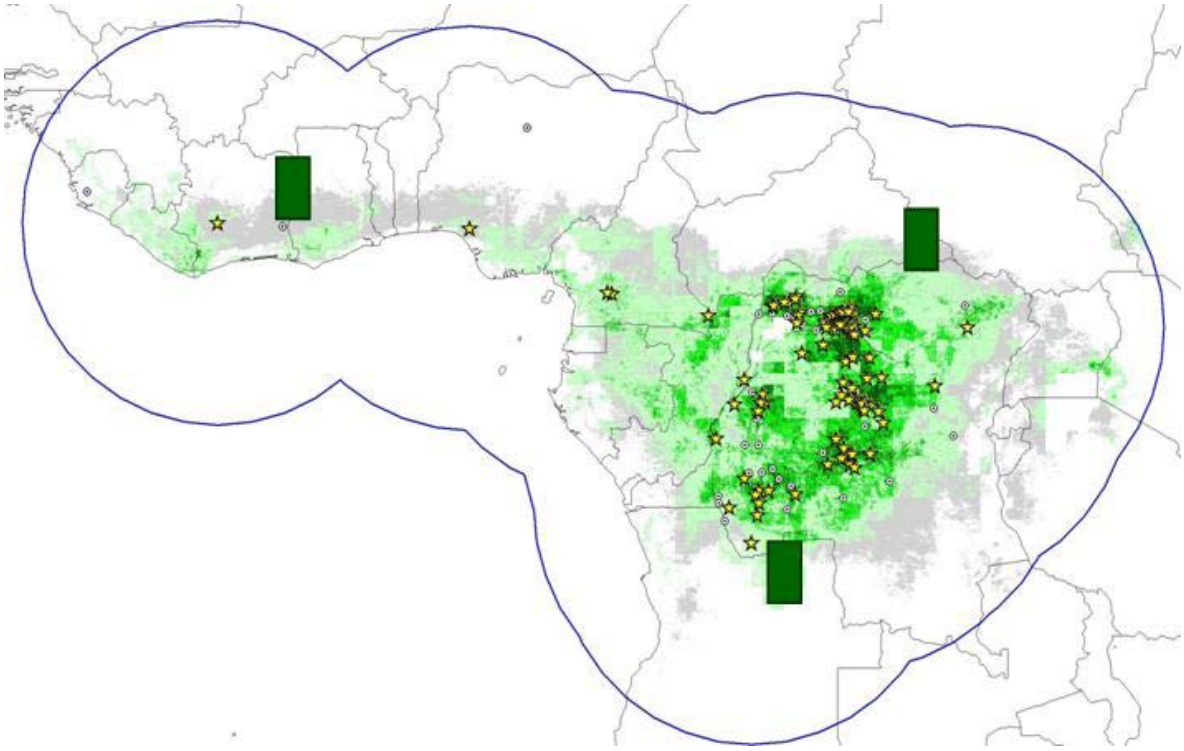
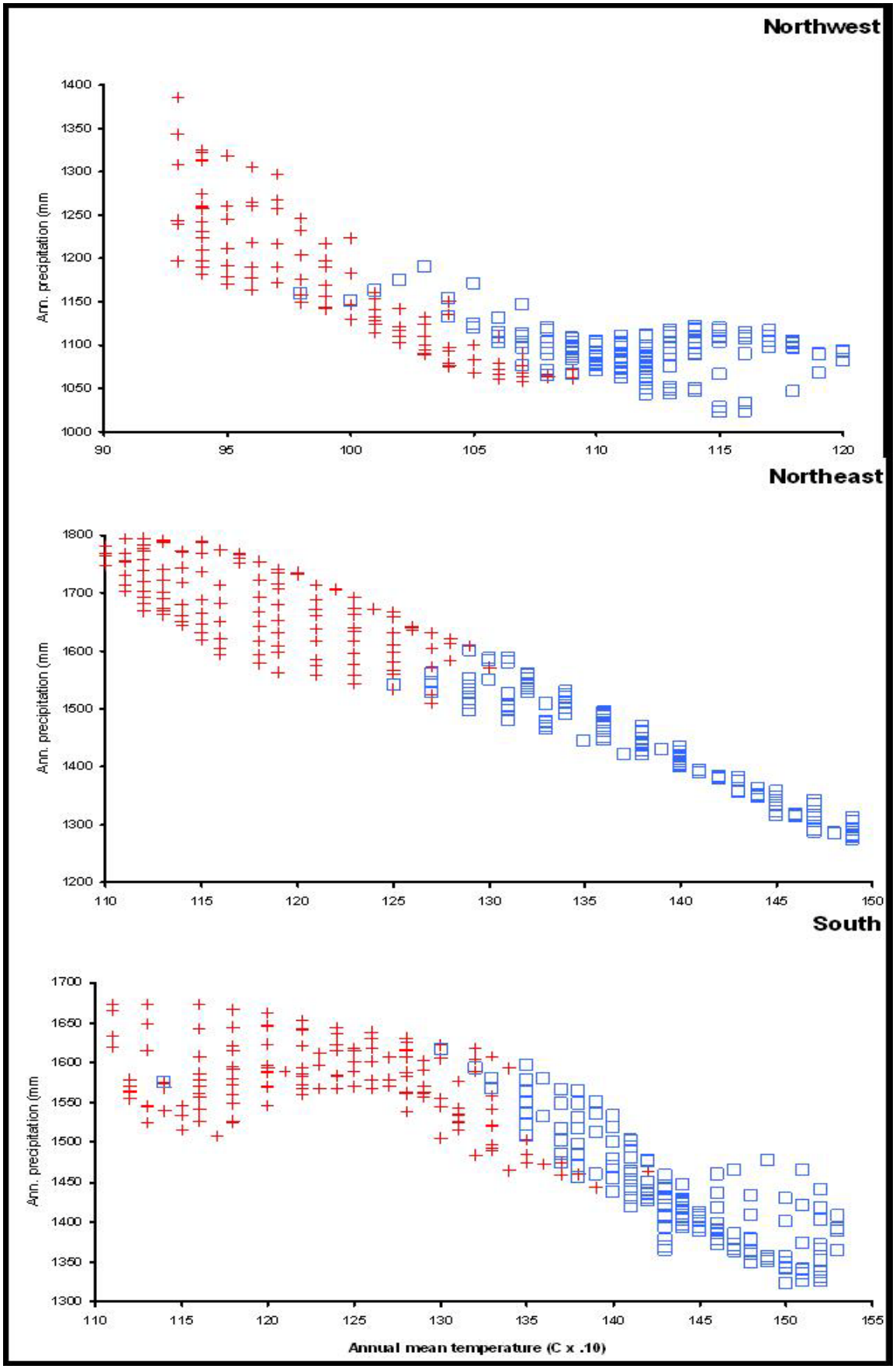


Figure 5: Sampling of environmental space for comparison of environmental characteristics. Each transect crosses different sectors of monkeypox range boundary and extends 50% into predicted favorable and unfavorable ecologic space.


Northwest
Ann. precipitation (mm
1400
1350
1300
1250
1200
1150
1100
1050
1000
90
95
100
105
110
115
120
Northeast
Ann. precipitation (mm
1800
1700
1600
1500
1400
1300
1200
110
115
120
125
130
135
140
145
150
South
Ann. precipitation (mm
1700
1650
1600
1550
1500
1450
1400
1350
1300
110
115
120
125
130
135
140
145
150
155
Annual mean temperature (C x .10)

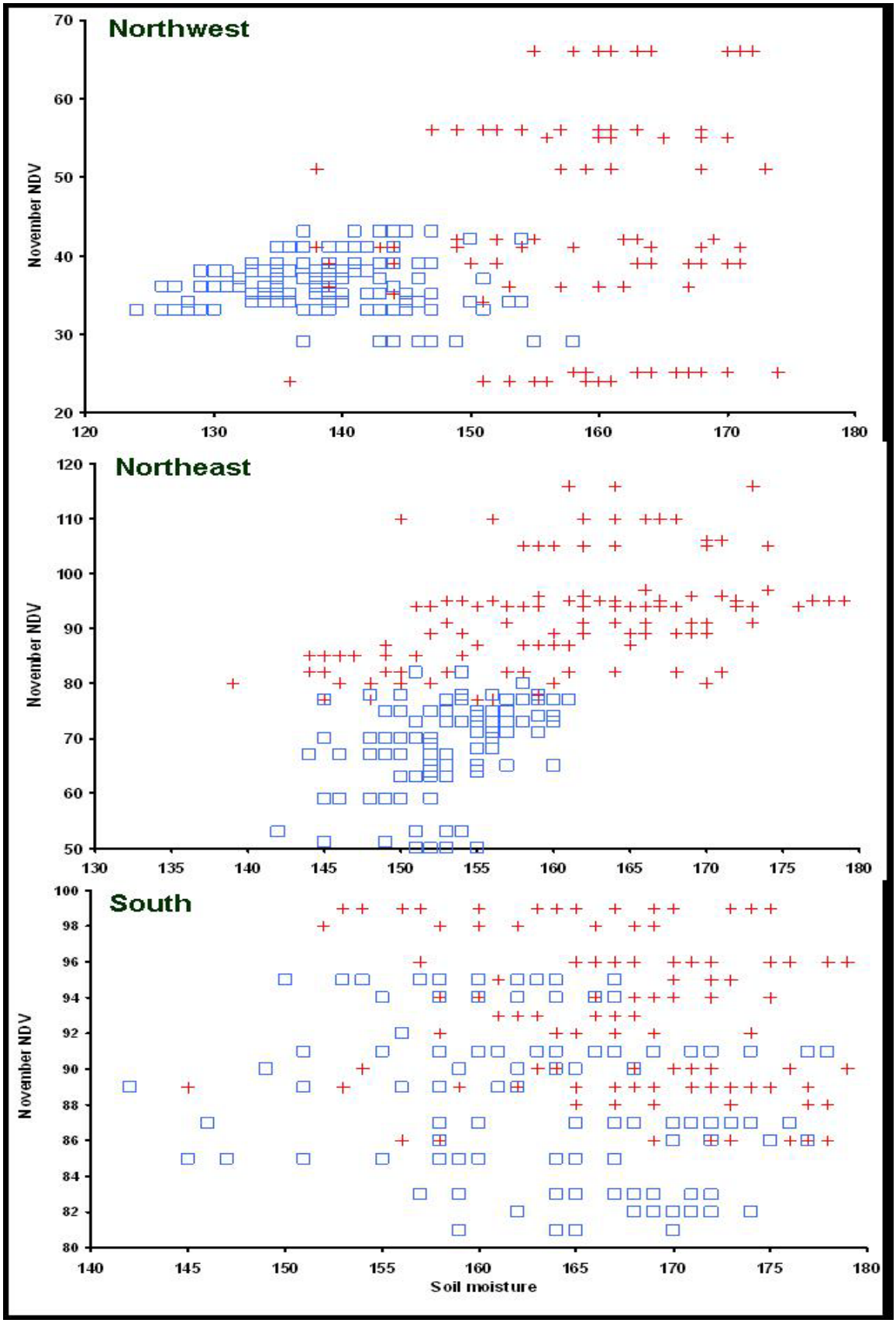

Northwest
November NDV
70
60
50
40
30
20
120
130
140
150
160
170
180
Northeast
November NDV
120
110
100
90
80
70
60
50
130
135
140
145
150
155
160
165
170
175
180
South
November NDV
100
98
96
94
92
90
88
86
84
82
80
140
145
150
155
160
165
170
175
180
Soil moisture

Figure 6: Summary of conditions of predicted presence versus absence within three transects crossing different sectors of the monkeypox range boundary (see Figure 5). Red plus signs indicate presence, blue open squares indicate absence. Predictions appear to be dependant on temperature and precipitation, whereas NDVI and soil moisture do not show consistent relationships in separating areas of presence and absence.

CHAPTER 4 - Discussion

Our goal in this study was to identify ecological factors relevant to understanding the geographic distribution of human monkeypox transmission in Africa using ecological niche modeling. Clearly, we were building upon and revisiting the results of a previous study via improved occurrence data georeferencing methodologies, and detailed exploration of more diverse environmental parameters (Levine et al. 2007). The ecological niche model produced, as in the previous analysis, identified areas of potential monkeypox distribution in Central and West Africa, focused in areas consistent with the known geographic distribution of humid lowland tropical forests (Levine et al. 2007).

In this study, favorable habitat was identified in 18 African countries, whereas in comparison, the previous study identified suitable habitat in 13 African countries. The most dramatic differences between the two studies were the inclusion of Angola, Benin, Brundi, Rwanda and Uganda, and the exclusion of Madagascar and Mozambique, from the our model. That the model developed in the present study is more realistic is perhaps supported by recent reports of an unknown orthopox virus in red colobus monkeys (*Piliocolobus* spp.) similar to other known orthopoxviruses in Uganda (Peterson and Nakazawa 2008). Distribution of favorable habitat is more homogenously distributed throughout West Africa in the model constructed in this study. A geographic break in the potential distribution of human monkeypox between West and Central Africa that may correspond to the distributional gap between the two monkeypox clades is present in both studies, however the geographic break identified in this study is more refined and corresponds specifically the Cameroon Range (also known as the Cameroon Highlands), a chain of mountains and volcanoes found in the border region of eastern Nigeria and western Cameroon.(Chen et al. 2005, Likos et al. 2005, Reed et al. 2004, Mackett and Archard 1979).

One likely reason for the differences between the two studies may be the bioclimatic variables used. This study made use of a more diverse suite of climatic variables reflecting annual, seasonal, and monthly patterns allowing a more refined view

of climate dimensions. What is more, the WorldClim data set is resolved spatially to 10', a 9-fold improvement in spatial resolution over the data sets used in the previous study (Hijmans et al. 2005, WorldClim). Finally, this study included aspects of land surface reflectance, soil, features, and vegetation characteristics, which offer additional information to our model by means of summarizing aspects of land cover—as such, descriptors such as net primary productivity; potential evapotranspiration; and soil carbon, pH, and moisture were included in our final models, and proved highly informative in model development.

Our ecological niche models are based on human case occurrence data collected by the CDC and WHO in 1970-1987. Biases in raw case data are well-known, including sampling bias, detection and reporting biases, and other factors that may distort the picture of the actual distribution of the species with respect to ecological and environmental factors (Araujo and Guisan 2006, Elith et al. 2006, Jones et al. 2008). In this respect, the niche modeling step employed in both studies – to some degree – allows a less-biased and more objective view of environmental distributions of species. Even given the niche modeling inferences, however, some adjustments must be made to determine which occurrence points are suitable for analysis (Peterson 2006, Peterson 2008a, Peterson 2008b). The case occurrence data set used in this study consisted of 404 laboratory-confirmed cases of human monkeypox in Africa, but only 127 were sufficiently precise for inclusion in the model. The previous analysis used the same case occurrence data set, and selected 156 occurrences for inclusion, but did not filter case occurrences based on spatial precision—as such, occurrences may have been included in the first data set that referred to broader regions or that were nebulous regarding precise location; this imprecision can produce overly broad estimates of ecological niches (Levine et al. 2007). The point-radius method that we have employed considers a "locality" as a geographic point combined with a radius that encompasses any associated uncertainties (Wieczorek et al. 2004). This approach has recently been recommended for broader application to reporting of disease occurrence (Peterson 2008b).

The previous study concluded that informative model layers included aspect, elevation, flow accumulation, flow direction, land cover, and topographic index (among others), that did not appear to contribute importantly to our models. Possible explanations

for these differences may be associated with data sources or spatial resolution, but most likely are a consequence of different statistical analyses used in the two studies. Although both studies utilized jackknife approaches, final statistical analyses in the previous study were performed using *t*-tests and the Kappa statistic, both of which are easily confounded by pseudoreplication of points (creating artificially large sample sizes), by prevalence of the phenomenon across the landscape, and by correlations among environmental variables (Fleiss 1971, Press et al. 1992, Viera and Garrett 2005).

Diseases and Niche Modeling

Incidences of infectious disease emergence appear to be increasing: from 1940-2004, ~ 335 newly emerging, or re-emerging infectious disease events occurred (Jones et al. 2008). During that period, zoonotic pathogens were responsible for the majority of emerging disease (60.3%), of which most (71.8%) were caused by pathogens of wildlife origin (e.g., Hantavirus, Nipah virus, Sudden Acute Respiratory Syndrome [SARS]) (Jones et al. 2008). From 1990-2000, the number of emerging disease events caused by wildlife pathogens increased by 52% over previous periods, and incidence of emerging disease caused by vector-borne pathogens also increased by 28.8% (Jones et al. 2008). The resulting scenario is one in which emerging zoonotic diseases will have significant impacts on local and global public health and economies (Jones et al. 2008).

The niche modeling methodologies demonstrated here may be used to summarize spatial patterns of disease transmission and risk, offering several advantages over commonly-used spatial and landscape epidemiology methodologies (Costa et al. 2002, Papes and Gaubert 2007, Peterson 2001, Peterson et al. 2002b, Peterson et al. 2002a, Peterson 2003, Peterson et al. 2006b, Ward 2007, Austin et al. 1990). The customary spatial analyses often identify broad trends, and as such are not fully applicable to characterizing fine details of disease transmission that may be highly dependent on local conditions. Because the spatial resolution of ecological niche models is limited only by the spatial precision of the occurrence data and environmental data, the resulting picture is much more refined (Peterson et al. 2006b, Peterson et al. 2007). What is more, niche modeling approaches are applicable even when sample sizes are relatively small, as demonstrated in recent analyses of the geography of Marburg virus transmission to

humans (Peterson et al. 2006a, Peterson 2007). Niche models also permit exploration of the potential geography and ecology of disease transmission even across novel landscapes (Peterson 2001, Peterson 2003, Peterson et al. 2006a).

Several issues must be addressed before ecological niche modeling methods can be applied fully to emerging disease and disease transmission systems. Identification and selection of key environmental datasets is particularly significant in building maximally accurate models: for example, climate data may provide longer temporal applicability, but remotely-sensed data provide a finer spatial resolution view of ecological landscapes (Peterson et al. 2006b). Analyzing these two data in resources tandem as we have done herein may offer advantages regarding identification of key environmental factors that could provide important insights into the transmission biology of diseases (Press et al. 1992).

The two-level jackknife analysis used in this paper offers a means of identifying environmental variables most informative for model development. Variables were first evaluated in suites to understand the significance of general classes of variables (Figure 2). The most significant suites included temperature and precipitation, as well as aspects of soils and surface reflectance. Individual variables were then assessed: the most informative climate variables identified included annual precipitation, annual mean temperature, maximum temperature of the warmest month, mean monthly diurnal temperature, minimum temperature of the coldest month, and precipitation of the driest month (Figure 3). Significant individual non-climatic variables included April NDVI, net primary productivity, November NDVI, potential evapotranspiration, soil moisture, and soil pH. Hence, our exploratory approach to environmental variable selection identified diverse informative variables, but still allowed reduction of dimensionality important to avoiding over-fitting (Sweeney et al. 2006).

Inspecting the NDVI variables through the year (Figure 3) suggests possible seasonal trends in signal. Specifically, NDVI variable importance is minimal in August, but increases rapidly through November, remains high through April, and then declines. What this trend suggests in terms of monkeypox natural history is unclear, given that little is known of its natural history; however, the trend is clear. Further field studies of

monkeypox may help to illuminate this association, or may in turn be illuminated by the insight.

Conclusions

The purpose of this study was to build upon and improve insights from a previous ecological niche modeling analysis of the potential distribution of monkeypox in Africa. Our results emphasize the importance of selecting the most appropriate and informative environmental data for ecological niche modeling. Models based on simple environmental data sets may be overly general in nature, lacking detail owing to broad interpolation and smoothing inherent in the process of generating the climate coverages (Nakazawa et al. 2007). Refinements such as filtering occurrence localities based on spatial precision can avoid imprecision resulting from uncertain geolocation (Wieczorek et al. 2004, Peterson 2008b). As ecological niche modeling continues to evolve as a tool in epidemiology and public health, focused studies evaluating these points may prove useful.

CHAPTER 5 - Bibliography

Agriculture, U. S. D. A. (2002) Agricultural bioterrorism protection act of 2002: listing of biological agents and toxins and requirements and procedures for notification and possession. *Federal Register,* 67**,** 383-389.

Araujo, M. B. & A. Guisan (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography,* 33**,** 1677-1688.

Arita, I., Z. Jezek, L. Khodakevich & K. Ruti (1985) Human monkeypox: A newly emerged orthopoxvirus zoonosis in the tropical rain forests of Africa. *American Journal of Tropical Medicine and Hygiene,* 34**,** 781-789.

Austin, M. P., A. D. Nicholls & C. R. Marguleis (1990) Measurement of the realized quantitative niche; Environmental niches of 5 Eucalyptus species. *Ecological Monographs,* 60**,** 161-177.

Breman, J. G. 2000. Monkeypox: an emerging infection for humans? In *Emerging Infections,* eds. W. Scheld & C. J. Hughes, 45-67. Washington D. C.: ASM Press.

Breman, J. G., R. Kalisa, M. V. Steniowski, E. Zanotto, A. Grumydo & I. Arita (1980) Human monkeypox, 1970-79. *Bulletin of the World Health Organization,* 58**,** 165-182.

Chen, N., G. Li, M. K. Liszewski, J. P. Atkinson, P. B. Jahrling, Z. Feng, J. Schriewer, C. Buck, C. Wang, E. J. Lefkowitz, J. J. Esposito, T. Harms, I. K. Damon, R. L. Roper, C. Upton & R. M. L. Buller (2005) Virulence differences between monkeypox virus isolates from West Africa and the Congo basin. *Virology,* 340**,** 46-63.

Costa, J., A. T. Peterson & C. B. Beard (2002) Ecological niche modeling and differentiation of populations of Triatoma brasiliensis nevia, 1911, the most important Chagas' Disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae). *The American Journal of Tropical Medicine and Hygiene,* 67**,** 516-520.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettman, J. R. Leathwick, A. Lehmann, J. L. Lucia, G. Lohman, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Sachelti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz & N. E. Zimmerman (2006) Novel methods improve prediction of species distributions from occurrence data. *Ecography,* 29**,** 129-151.

Esposito, J. J. & J. C. Knight (1985) Orthopoxvirus DNA: a comparison of restriction profiles and maps. *Virology,* 143**,** 230-251.

Fleiss, J. L. (1971) Measuring nominal scale agreement among many raters. *Psychological Bulletin,* 76**,** 378-382.

Foster, S., E. W. Brind, D. L. Hutchins, J. M. Pifer, B. Lourie, C. R. Moser, E. C. Cummings, O. E. K. Kutey, R. E. Eke, J. B. Titus, E. A. Smith, J. W. Hicks & W.

H. Foege (1972) Human monkeypox. *Bulletin of the World Health Organization,* 46**,** 569-576.
Grinnell, J. (1917) Field tests of theories concerning distributional control. *The American Naturalist,* 51**,** 115-128.
--- (1924) Geography and Evolution. *Ecology,* 5**,** 225-229.
Hernandez, P. A., C. H. Graham, L. L. Master & D. L. Albert (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography,* 29**,** 773-785.
Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones & A. Jarvis (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* 25**,** 1965-1978.
Huhn, G. D., A. M. Bauer, K. Yorita, M. B. Graham, J. Sejvar, A. Likos, I. K. Damon, M. G. Reynolds & M. J. Kuehnert (2005) Clinical characteristics of human monkeypox, and risk factors for severe disease. *Clinical Infectious Diseases,* 41**,** 1742-1751.
Hutin, Y. J., R. J. Williams, P. Malfait, R. Peabody, V. N. Loparev, S. L. Rupp, M. Rodriguez, J. C. Knight, F. K. Tshioko, A. S. Khan, M. V. Szczeniowski & J. J. Esposito (2001) Outbreak of human monkeypox, Democratic Republic of Congo, 1996-1997. *Emerging Infectious Diseases,* 7**,** 434-438.
Janseghers, L., M. Matamba, J. Colarert, J. Vandepitte & J. Desmyter (1984) Fatal monkeypox in a child in Kikwit, Zaire. *Annales de la Societe belge de Medicine Tropicale,* 64**,** 295-298.
Jezek, A., A. L. Gromyko & M. V. Szczeniowski (1983) Human monkeypox. *Journal of Hygiene, Epidemiology, Microbiology, and Immunology,* 27**,** 13-28.
Jezek, A., S. S. Marennikova, M. Mutumbo, J. H. Nakano, K. M. Paluku & M. Szczeniowski (1986) Human monkeypox: a study of 2510 contacts of 214 patients. *Journal of Infectious Diseases,* 154**,** 551-555.
Jezek, Z., B. Grab, M. V. Szczeniowski, K. M. Paluku & M. Mubuto (1988) Clinico-epidemiological features of monkeypox patients with an animal or human source of infection. *Bulletine of the World Health Organization,* 66**,** 459-464.
Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman & P. Daszak (2008) Global trends in emerging infectious diseases. *Nature,* 451**,** 990-994.
Khodakevich, L., Z. Jezek & K. Kinzanzka (1986) Isolation of monkeypox virus from wild squirrel infected in nature. *The Lancet,* 1**,** 98-99.
Khodakevich, L., Z. Jezek & D. Messinger (1988) Monkeypox virus: Ecology and public health significance. *The Bulletin of the World Health Organization,* 66**,** 747-752.
Ladnyj, I. D., P. Ziegler & E. Kima (1972) A human infection causes by monkeypox virus in Basankusu Territory, Democratic Republic of the Congo. *The Bulletin of the World Health Organization,* 46**,** 593-597.
Learned, L. A., M. G. Reynolds, D. Wassa Wassa, Y. Li, V. A. Olson, K. Karem, L. L. Stempora, Z. H. Braden, R. Kline, A. Likos, F. Libama, H. Moudzeo, J. D. Bolanda, P. Tarangonia, P. Boumandoki, P. Formenty, J. M. Harvey & I. K. Damon (2005) Extended interhuman transmission of monkeypox in a hospital community in the Republic of the Congo, 2003. *American Journal of Tropical Medicine and Hygiene,* 73**,** 428-434.

Levine, R. S., A. T. Peterson, K. L. Yorita, D. Carroll, I. K. Damon & M. G. Reynolds (2007) Ecological niche and geographic distribution of human monkeypox in Africa. *PlosOne,* 1**,** 1-7.

Lewis-Jones, S. (2004) Zoonotic poxvirus infections in humans. *Current Opinion in Infectious Disease,* 17**,** 81-89.

Likos, A. M., S. A. Sammons, V. A. Olson, A. M. Frace, M. Olsen-Rasmussen, W. Davidson, R. Galloway, M. L. Khristova, M. G. Reynolds, H. Zhao, D. S. Carroll, A. Curns, P. Formenty, J. J. Esposito, R. L. Regnery & I. K. Damon (2005) A tale of two clades: monkeypox virus. *Journal of General Virology,* 86**,** 2661-2672.

Mackett, M. & L. Archard (1979) Conservation and variation in orthopoxvirus genome structure. *Journal of General Virology,* 45**,** 683-701.

Meyer, A., J. J. Esposito, F. Gras, T. Kolakowski, M. Fatras & G. Muller (1991) First appearance of monkeypox in human beings in Gabon. *Medicine Tropicale,* 51**,** 53-57.

Nakazawa, Y., R. Williams, A. T. Peterson, P. Mead, E. Staples & K. Gage (2007) Climate change effects on plague and tularemia in the United States. *Vector Borne and Zoonotic Disease,* 7**,** 529-540.

Nalca, A., A. W. Rimoin, S. Bavari & C. A. Whitehouse (2005) Reemergence of monkeypox: Prevalence, diagnostics, and countermeasures. *Clinical Infectious Diseases,* 41**,** 1765-1771.

Papes, M. & P. Gaubert (2007) Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity and Distributions,* 13**,** 890-902.

Parker, S., A. Nuara, R. M. Buller & D. A. Schultz (2007) Human monkeypox: An emerging zoonotic disease. *Future Microbiology,* 2**,** 17-34.

Peterson, A. T. (2001) Predicting species geographic distributions based on ecological niche modeling. *Condor,* 103**,** 599-605.

--- (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology,* 78**,** 419-433.

--- (2006) Ecologic niche modeling and spatial patterns of disease transmission. *Emerging Infectious Diseases,* 12**,** 1822-1826.

--- (2007) Ecological niche modelling and understanding the geography of disease transmission. *Veterinaria Italiana,* 43**,** 393-400.

--- (2008a) Biogeography of diseases: A framework for analysis. *Naturwissenschaften,* 95**,** 483-491.

--- (2008b) Improving methods for reporting spatial epidemiologic data. *Emerging Infectious Diseases***,** 1335-1337.

Peterson, A. T., L. G. Ball & K. P. Cohoon (2002a) Predicting distributions of Mexican birds using ecological niche modeling methods. *Ibis,* 144**,** e27-e32.

Peterson, A. T., D. Carroll & M. J. N. (2004) Potential mammalian filovirus reservoirs. *Emerging Infectious Diseases,* 10**,** 2073-2081.

Peterson, A. T. & K. C. Cohoon (1999) Sensitivity of distribution prediction algorithms to geographic completeness. *Ecological Modeling,* 117**,** 159-164.

Peterson, A. T., R. R. Lash, D. S. Carroll & K. M. Johnson (2006a) Geographic potential for outbreaks of Marburg hemorrhagic fever. *The American Journal of Tropical Medicine and Hygiene,* 75**,** 9.

Peterson, A. T., C. Martinez-Compos, Y. Nakazawa & E. Martinez-Meyer (2005) Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 99**,** 647.

Peterson, A. T. & Y. Nakazawa (2008) Environmental data sets matter in ecological niche modelling: an example with Solenopsis invicta and Solenopsis richteri. *Global Ecology and Biogeography,* 17**,** 135-144.

Peterson, A. T., M. Papes & M. Eaton (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography,* 30**,** 550-560.

Peterson, A. T., V. Sanchez-Cordero, E. Martinez-Meyer & A. G. Navarro-Siguenza (2006b) Tracking population extirpations via melding ecological niche modeling with land-cover information. *Ecological Modeling,* 195**,** 229-236.

Peterson, A. T., D. R. Stockwell & D. A. Kluza. 2002b. Distributional prediction based on ecological niche modeling of primary occurrence data. In *Predicting species occurrences: Issues of scale and accuracy,* eds. J. M. Scott, P. J. Heglund, M. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall & F. B. Samson, 617-623. Washington, D. C.: Island Press.

Phillips, S. J., R. P. Anderson & R. E. Schapire (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modeling,* 190**,** 231-259.

Phillips, S. J., M. Dudik & R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. In *21st International conference on machine learning,* 655-662. New York, NY: AMC Press.

Press, W. H., B. P. Flannery, S. A. Tuekolsky & W. T. Vetterling. 1992. *Numerical recipes in C: The art of scientific computing*. Cambridge, MA: Cambridge University Press.

Reed, K., D., J. W. Melski, M. B. Graham, R. L. Regnery, M. J. Sotir, M. V. Wegner, J. J. Kazmierezak, E. J. Stratman, Y. Li, J. A. Fairley, G. R. Swain, V. A. Olson, E. K. Sargent, S. C. Kehl, F. M. A., R. Kline, S. L. Foldy, J. P. Davis & D. I. K. (2004) The detection of monkeypox in humans in the western hemisphere. *The New England Journal of Medicine,* 350**,** 342-350.

Reynolds, M. G., K. L. Yorita, M. J. Kuehnert, W. B. Davidson, G. D. Huhn, R. C. Holman & I. K. Damon (2006) Clinical manifestations of human monkeypox influenced by route of infection. *The Journal of Infectious Diseases,* 194**,** 773-80.

Sale, T. A., J. W. Melski & E. J. Stratmen (2006) Monkeypox; and epidemiologic and clinical comparison of African and US disease. *Journal of the American Academy of Dermatology,* 55**,** 478-481.

Santiago, R. R. (2005) Predicting the distribution of the amphibian pathogen Batrachochytrium dendrobatidis in the New World. *Biotropica,* 37**,** 209-221.

Sbrana, E., S.-Y. Xiao, P. C. Newman & R. B. Tesh (2007) Comparative pathology of North American and Central African strains of monkeypox virus in a ground squirrel model of the disease. *American Journal of Tropical Medicine and Hygiene,* 76**,** 155-164.

Stanford, M. M., G. McFadden, G. Karupiah & G. Chaudhri (2007) Immunopathogenesis of poxvirus infections: Forecasting the impending storm. *Immunology and Cell Biology,* 85**,** 93.

Stockwell, D. & A. T. Peterson (2003) Comparison of resolution of methods used in mapping biodiversity patterns from point-occurrence data. *Ecological Indicators,* 3**,** 213-221.

Sweeney, A. W., N. W. Beebe, R. D. Cooper, J. T. Bauer & A. T. Peterson (2006) Environmental factors associated with distribution and range limits of malaria vector Anopheles farauti in Australia. *Journal of Medical Entomology,* 43**,** 1068-1075.

Townsend, P. A., V. Sanchez-Cordera & E. Martinez-Meyer (2005) Tracking population extirpations via melding ecological niche modeling with land-cover information. *Ecological Modelling,* 195**,** 229-236.

Viera, A. J. & J. M. Garrett (2005) Understanding interobserver agreement: the kappa statistic. *Family Medicine,* 37**,** 360-363.

Ward, D. F. (2007) Modeling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions,* 9**,** 723-725.

Wieczorek, J., Q. Guo & R. J. Hijmans (2004) The point radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science,* 18**,** 745-767.

WorldClim. www.worldclim.org.