A Journal of Macroecology

Global Ecology and Biogeography

RESEARCH PAPER

Using MODIS satellite imagery to predict hantavirus risk

Lina Cao[1]*, Thomas J. Cova[1], Philip E. Dennison[1] and M. Denise Dearing[2]

[1]*Department of Geography, University of Utah, Salt Lake City, UT 84112, USA,* [2]*Department of Biology, University of Utah, Salt Lake City, UT 84112, USA*

ABSTRACT

Aims Sin Nombre virus (SNV), a strain of hantavirus, causes hantavirus pulmonary syndrome (HPS) in humans, a deadly disease with high mortality rate (> 50%). The primary virus host is the deer mouse, and greater abundance of deer mice has been shown to increase the human risk of HPS. Our aim is to identify and compare vegetation indices and associated time lags for predicting hantavirus risk using remotely sensed imagery.

Location Utah, USA.

Methods A 5-year time-series of moderate-resolution imaging spectroradiometer (MODIS) satellite imagery and corresponding field data was utilized to compare various vegetation indices that measure productivity with the goal of indirectly estimating mouse abundance and SNV prevalence. Relationships between the vegetation indices and deer mouse density, SNV prevalence and the number of infected deer mice at various time lags were examined to assess which indices and associated time lags might be valuable in predicting SNV outbreaks.

Results The results reveal varying levels of positive correlation between the vegetation indices and deer mouse density as well as the number of infected deer mice. Among the vegetation indices, the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) produced the highest correlations with deer mouse density and the number of infected deer mice using a time lag of 1.0 to 1.3 years for May and June imagery.

Main conclusions This study demonstrates the potential for using MODIS time-series satellite imagery in estimating deer mouse abundance and predicting hantavirus risk. The 1-year time lag provides a great opportunity to apply satellite imagery to predict upcoming SNV outbreaks, allowing preventive strategies to be adopted. Analysis of different predictive indices and lags could also be valuable in identifying the time windows for data collection for practical uses in monitoring rodent abundance and subsequent disease risk to humans.

Keywords

Deer mice, hantavirus, MODIS, Sin Nombre virus, time-series, Utah, vegetation indices.

*Correspondence: Lina Cao, 260 S. Central Campus Dr., Rm. 270, Salt Lake City, Utah 84112, USA.
E-mail: lina.cao@geog.utah.edu

INTRODUCTION

In May 1993, an outbreak of hantavirus pulmonary syndrome (HPS) occurred among previously healthy young people in the Four Corners region of the south-western USA. HPS has a relatively high mortality rate (> 50%) and is characterized by acute respiratory distress (Nichol *et al.*, 1993; Glass *et al.*, 2000). HPS was traced to infection with Sin Nombre virus (SNV), a strain of hantavirus of which the deer mouse (*Peromyscus maniculatus*) is the primary reservoir (CDC, 1993; Childs *et al.*, 1994; Hjelle *et al.*, 1996). There is currently no vaccine or effective drug to prevent or treat HPS (Custer *et al.*, 2003; Buceta *et al.*, 2004), and for this reason there is a need to understand the nature of the virus as well as its spatial and temporal dynamics in order to

DOI: 10.1111/j.1466-8238.2010.00630.x
www.blackwellpublishing.com/geb

predict the risk of the disease and design effective prevention policies.

Environmental conditions, such as climate (Glass *et al.*, 2000; Yates *et al.*, 2002), seasonality (Cantoni *et al.*, 2001; Dearing *et al.*, 2009) and vegetation type (Boone *et al.*, 2000), have been associated with the geographic distribution of SNV in deer mouse populations in past studies. One of the main drivers for the inter-annual changes in HPS cases is thought to be fluctuation in precipitation and temperature. The El Niño of 1991–92 is believed to be the major climatic factor leading to the outbreak of HPS in 1993 in the south-western USA. The dramatic increase in rainfall is believed to have resulted in more food and thus increased local rodent populations such as deer mice (Parmenter *et al.*, 1993; Engelthaler *et al.*, 1999). This increase in deer mouse abundance enhances the human risk of HPS (Parmenter *et al.*, 1993). Childs *et al.* (1995) conducted a household-based, case–control study of environmental factors associated with HPS and found that a higher number of captured, infected deer mice is associated with the occurrence of HPS.

Yates *et al.* (2002) outlined the effects of climate change on the abundance of rodent populations in the trophic cascade hypothesis. As precipitation increases and temperature is more tolerable, net primary productivity increases and subsequently results in a larger deer mouse population. Increased mouse density increases viral transmission among mice and results in a larger number of dispersing mice. This leads to increased disease transmission to humans who come into contact with these mice. When environmental conditions become more severe mouse populations decline, but the environmental conditions at some locations allow survival of a sufficiently large rodent population for the virus to persist. Thus, it is important to identify the environmental conditions needed to maintain host populations of a sufficient size to sustain the virus. Hantaviruses are horizontally transmitted among members of the rodent population but are not vertically transmitted to offspring. They need a large host population to avoid local extinction (Glass *et al.*, 2007).

Satellite imagery has demonstrated value in linking environmental conditions to disease distribution and dynamics. For example, Linthicum *et al.* (1999) found that outbreaks of Rift Valley fever could be predicted up to 5 months in advance in Kenya using the normalized difference vegetation index (NDVI) and Pacific and Indian Ocean sea-surface temperature anomalies. Thomson *et al.* (1997) used meteorological satellite data to model the spatial and seasonal dynamics of infectious disease transmission and developed affordable early warning systems for malaria. A study of malaria prevalence in children in The Gambia used NDVI to measure changes in vegetation growth as proxy ecological variables representing changes in rainfall and humidity to predict the length and intensity of malaria transmission (Thomson *et al.*, 1999).

A few previous studies have utilized Landsat Thematic Mapper (TM) satellite imagery to study hantavirus dynamics. Glass *et al.* (2002) developed logistic regression models to predict risk to humans of HPS using Landsat TM imagery. They found that heavy rainfall associated with the El Niño–Southern Oscillation increased the rodent population and preceded HPS cases in the south-western USA (Glass *et al.*, 2000). Goodin *et al.* (2006) evaluated the relationship between land cover and hantavirus prevalence in rodents. Their land-cover map was derived from a variety of coarse-resolution satellite imagery depending on the type of land cover being mapped (e.g. along track scanning radiometer (ATSR) and SPOT-VGT). A positive relationship was found between agricultural land-cover disturbance and hantavirus in rodents in Paraguay. However, single-date satellite imagery can be limited in capturing the vegetation dynamics that affect rodent population dynamics.

The moderate resolution imaging spectroradiometer (MODIS), launched on the NASA satellites Terra (December 1999) and Aqua (May 2002), significantly improved the availability of data for epidemiological studies. There have been a few studies that utilized MODIS time-series data in modelling rodent pathogen transmission and predicting disease risk to humans. Glass *et al.* (2007) applied MODIS NDVI data to compare vegetation growth patterns in years of severe drought from 2002 to 2004 and found that high-risk HPS areas had higher levels of green vegetation and longer durations of greenness. Marston *et al.* (2007) modelled the spatial distribution of the rodent species that were the hosts of a parasitic tapeworm. The rodent distribution was modelled with landscape characteristics using four different types of remotely sensed data. Their results showed that the MODIS time-series image data provided the strongest relationships and explained the highest percentage deviance of the relationships present (up to 41.4%). These results support the idea that using time-series NDVI data can offer improved results over single-date imagery.

In this study we investigated the value of MODIS data for estimating rodent abundance and SNV prevalence with the goal of predicting hantavirus risk. We examined relationships between environmental conditions (vegetation greenness and moisture), deer mouse density and SNV prevalence using high-temporal-resolution (16 days) MODIS satellite imagery and multi-year field survey data. The central hypothesis is that vegetation indices can serve as proxies for deer mouse food availability that affects deer mouse abundance and SNV prevalence. We applied MODIS time-series imagery to measure vegetation productivity and to indirectly estimate mouse abundance and SNV prevalence in deer mouse populations.

METHODS

Study area

The study area is located in proximity to the Little Sahara recreation area in Juab County, Utah (39°40′ N, 112°15′ W). The elevation in this area ranges from 1600 to 1900 m. Dominant vegetation species in the area include Great Basin sagebrush (*Artemisia tridentata*) and Utah juniper (*Juniperus osteosperma*). Areas with high sagebrush cover are positively correlated with deer mouse abundance (Pearce-Duvet *et al.*, 2006). The preferred habitat for deer mice is sagebrush. This particular area has experienced heavy recreational use, mainly by all-terrain vehicles (ATVs) (Lehmer *et al.*, 2008). Landscape disturbance

due to recreational ATV use has created new roads, trails and open spaces. Trapping webs at the study area were selected based on low and high levels of disturbance. Low disturbance refers to those areas where ATVs have had little or no effect, and high disturbance refers to the areas with trails and vast expanses of bare ground caused by camping and the heavy use of recreational vehicles.

Field data

Rodents were live-trapped in the spring and autumn of 2004, 2005 and 2006. Using a web-based approach (Anderson *et al.*, 1983), 12 3.14-ha trapping sites were established for the spring and autumn of 2004, 2005 and 2006. There have been two general field sampling designs for estimating small-mammal populations: grid-based regimes and web-based approaches (typically transect lines or trapping webs) (Burnham *et al.*, 1980; Anderson *et al.*, 1983; Buckland *et al.*, 1993, 2001). Web-based trapping has been utilized in a number of studies, including small-mammal studies at a number of sites within the US Long Term Ecological Research Network, monitoring programmes for rodent-borne zoonotic diseases conducted by public health agencies in the USA (Parmenter *et al.*, 1999), and global-change/biodiversity programmes around the world. The trapping web design in this study is the most commonly used trapping arrangement in longitudinal studies of hantavirus (Abbott *et al.*, 1999; Calisher *et al.*, 1999; Kuenzi *et al.*, 1999; Mills *et al.*, 1999a,b; Root *et al.*, 1999).

Trapping webs are centred on a reference location with traps emanating from this point. The trapping web contains 12 100-m trap transects radiated from the central trap at angles of 30° from one another. Transects were numbered from 1 to 12, with the first transect facing south and other transects incremented in a clock-wise direction. Twelve traps are placed on each trap transect. For each trap transect, the first four traps are spaced 5 m apart and the other eight traps are 10 m apart. Each trap is coded and marked by the transect number and its location along each transect labelled by letters from A to L with A for the closest trap to the centre and L for the furthest.

In each sampling period, traps were set for each trapping web for three consecutive nights. Small-mammal processing was conducted according to protocols and methods for trapping and sampling small mammals for virological testing (Mills *et al.*, 1995). Ear tags with individual numbers were attached to newly captured mice. Captured mice were weighed and recorded for trap code, sex, reproductive status, weight, ear tag number, wounds, scars, etc. Blood was drawn from deer mice for testing for the presence of SNV. Rodents were released where they were captured.

Enzyme-linked immunosorbent assay for detection of Sin Nombre virus antibody was performed on all blood samples (Feldman *et al.*, 1993; Lehmer *et al.*, 2008). Prevalence is reported as the percentage of infected deer mice at each sampling site. For this reason estimates of prevalence increase in accuracy with increasing sample size.

MODIS satellite imagery

This study utilized MODIS surface reflectance 16-day composite data with 500 m spatial resolution to derive environmental conditions of the study sites across multiple years (2003–06). MODIS instruments capture data in 36 spectral bands ranging in wavelength from 0.4 to 14.4 μm and at varying spatial resolutions (2 bands at 250 m, 5 bands at 500 m and 29 bands at 1 km). They provide twice-a-day global coverage at 250 m (red, near-infrared), 500 m (mid-infrared) and 1000 m resolution (thermal infrared). The datasets are available to the public *c.* 1 week after acquisition, and they are designed to provide measurements of large-scale global dynamics and processes occurring on the land and in the oceans and lower atmosphere. MODIS imagery has been successfully used for a variety of applications including quantifying vegetation cycles, assessing land-cover change and mapping the spatial distribution of habitats (Jin & Sader, 2005; Beck *et al.*, 2006; Lunetta *et al.*, 2006; Xiao *et al.*, 2006).

Temporal compositing methods

Temporal compositing of remote sensing time-series data is a common practice to compress data and reduce the impacts of cloud effects and changing view geometry (Qi & Kerr, 1997). Temporal compositing algorithms analyse the pixel values across time and select the single best pixel value to represent the entire time period (Dennison *et al.*, 2007).

Many temporal compositing algorithms have been developed for moderate- to coarse-resolution remote sensing systems. Dennison *et al.* (2007) introduced a new class of compositing algorithms based on two measures of spectral similarity, end-member average root mean square error (EAR) and minimum average spectral angle (MASA). Their research demonstrated that these novel algorithms reduce short-term variability in spectral indices across several land-cover types (Dennison *et al.*, 2007). A set of 16-day MODIS compositing data across March to July for 5 years (2002–06) using MASA and EAR methods was created and used to derive several vegetation indices.

Vegetation indices from MODIS

The correlations between a set of vegetation indices calculated from 5 years of MODIS data and deer mouse density, hantavirus prevalence and the number of infected mice for each site across 4 years (2003–06) were examined using linear regression. The sample sizes varied from 36 for spring to 43 for autumn because a few additional sites were added in autumn 2003 and 2004. Four commonly used vegetation indices were calculated, including NDVI (Rouse *et al.*, 1973), the enhanced vegetation index (EVI; Huete *et al.*, 2002), normalized difference water index (NDWI; Gao, 1996) and visible atmospherically resistant index (VARI; Gitelson *et al.*, 2002). NDVI, EVI and VARI are greenness indices based on chlorophyll absorption and near-infrared reflectance and/or visible reflectance, while NDWI is a moisture index based on near-infrared water absorption.

NDVI is a very simple, well-known and widely used remote sensing vegetation index. It is calculated from the individual measurements as follows:

$$\mathrm{NDVI} = (\rho_{\mathrm{NIR}} - \rho_{\mathrm{RED}})/(\rho_{\mathrm{NIR}} + \rho_{\mathrm{RED}}) \quad (1)$$

where ρ_{RED} and ρ_{NIR} stand for the red and near-infrared reflectance. The resulting values ranges from −1.0 to +1.0.

EVI was developed to improve the NDVI by accounting for soil and atmospheric interference (Justice *et al.*, 1998; Huete *et al.*, 1997, 2002). EVI normalizes the red band reflectance by the blue band reflectance (ρ_{Blue}) (Huete *et al.*, 1997). EVI is more sensitive in vegetation with a high leaf area index (LAI), where NDVI saturates quickly and shows very little dynamic range for high-LAI canopies in crop fields (Boegh *et al.*, 2002). EVI is formulated as:

$$\mathrm{EVI} = G(\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}})/(\rho_{\mathrm{NIR}} + \mathrm{C1}\,\rho_{\mathrm{Red}} - \mathrm{C}\,2\rho_{\mathrm{Blue}} + L) \quad (2)$$

where $G = 2.5$, C1 = 6, C2 = 7.5 and $L = 1$. VARI is a vegetation index based entirely on visible reflectance (ρ_{Green}, ρ_{Red} and ρ_{Blue}) (Gitelson *et al.*, 2002). VARI is calculated as:

$$\mathrm{VARI} = (\rho_{\mathrm{Green}} - \rho_{\mathrm{Red}})/(\rho_{\mathrm{Green}} + \rho_{\mathrm{Red}} - \rho_{\mathrm{Blue}}). \quad (3)$$

NDWI was introduced by Gao (1996) to assess water content using near-infrared water absorption. NDWI increases with vegetation water content. It is defined as follows:

$$\mathrm{NDWI} = (\rho_{0.86\,\mu\mathrm{m}} - \rho_{1.24\,\mu\mathrm{m}})/(\rho_{0.86\,\mu\mathrm{m}} + \rho_{1.24\,\mu\mathrm{m}}). \quad (4)$$

The 0.86 μm and 1.24 μm channels of MODIS ($\rho_{0.86\,\mu\mathrm{m}}$ and $\rho_{1.24\,\mu\mathrm{m}}$) are band 2 and band 5, respectively.

Time lags

Glass *et al.* (2002) found that there was an apparent 1-year lag between the end of the 1997–98 El Niño and the increase in hantavirus prevalence in high-risk areas, which was also the case in the 1993 HPS outbreak preceded by the 1991–92 El Niño event. This time lag is also in line with the trophic cascade hypothesis (Yates *et al.*, 2002).

Vegetation indices were calculated for MODIS composites from March to July, 2002–06. This time period coincides with vegetation green-up in the study area, and the index values during this time period should indicate vegetation productivity. Time-lag effects between the spring vegetation indices (2002–06) and deer mouse density and hantavirus prevalence (2003–06) were explored with lags of 0, 0.3, 1.0 and 1.3 years. A 0-year lag examined the correlation between spring vegetation indices and the same-year spring field data. The 0.3-year lag examined the correlation between spring vegetation indices and autumn field data in the same year. The 1-year time lag examined the correlation between spring vegetation indices and field data from the following spring. The 1.3-year time lag examined correlations between spring vegetation indices and autumn field data in the following year. Linear regression was used and r^2 was calculated and compared.

RESULTS

Deer mouse density and vegetation indices

Correlations between vegetation indices and deer mouse density typically peaked in May or June for NDVI and EVI and in March for NDWI. For VARI, the correlation with density peaked in March at 0- and 0.3-year time lags and was highest in May at 1.0- and 1.3-year lags. Figure 1 shows an example of changing correlations between EVI and deer mouse density for each composite period, and at the four different lag times. At peak r^2, linear correlations between deer mouse density and NDVI/EVI at a 1.3-year lag were mostly significant. Figure 2 shows an example of the relationship between 10 June EVI and density at a 1.0-year time lag.

The maximum r^2 for each vegetation index and lag are shown in Table 1. Linear regression showed that all vegetation indices had significant ($P < 0.01$) correlations with deer mouse density at various time lags. For correlations between deer mouse density and NDVI/EVI over various time lags, the 1.0- and 1.3-year lags yielded higher r^2 than the 0- and 0.3-year lags, where the 1.3-year lag had the highest r^2. Correlations were highest for NDVI at a 1.3-year lag, with an r^2 of 0.50 and significance greater than 99.9%. The intercept and slope coefficients of the linear regression were −99.39 and 527.26. Correlations were also high for EVI at 1.0- and 1.3-year lags, with r^2 of 0.46 and 0.47 and significance greater than 99.9%. The intercepts and slope coefficients of the regression lines were −77.68 and 772.30 for a 1.0-year lag and −100.15 and 893.44 for a 1.3-year lag, respectively. Significant correlations were also found between VARI and deer mouse density at all time lags,

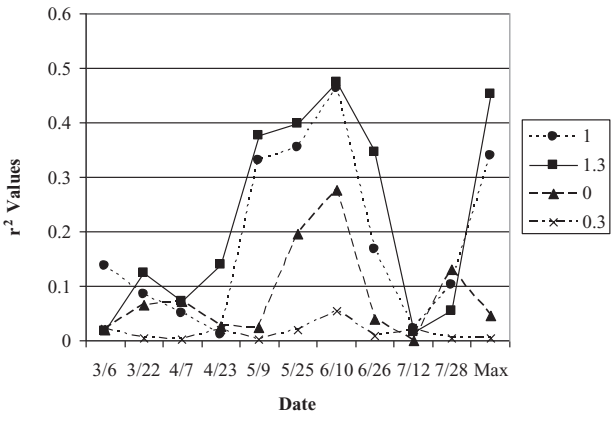

Figure 1 The r^2 of deer mouse density and the enhanced vegetation index (EVI) between March and July at 0-, 0.3-, 1.0- and 1.3-year lags. The correlation peaked in June, and the 1.0- and 1.3-year lag yielded the highest significant r^2 (0.46 and 0.47, respectively).

Table 1 Maximum r^2 values and dates of maximum r^2 values between March and July for normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), visible atmospherically resistant index (VARI) and normalized difference water index (NDWI) regressed against the deer mouse density at 0-, 0.3-, 1.0- and 1.3-year lags.

Time lag (years)	Maximum NDVI r^2	Date of maximum NDVI r^2	Maximum EVI r^2	Date of maximum EVI r^2	Maximum VARI r^2	Date of maximum VARI r^2	Maximum NDWI r^2	Date of maximum NDWI r^2
0	0.31*	6 March	0.28*	10 June	0.29*	6 March	0.35*	6 March
0.3	0.18	6 March	0.05	10 June	0.21*	6 March	0.23*	6 March
1	0.32*	9 May	0.46*	10 June	0.24*	9 May	0.11	25 May
1.3	0.50*	9 May	0.47*	10 June	0.27*	9 May	0.07	25 May

*Significance ($P < 0.01$).

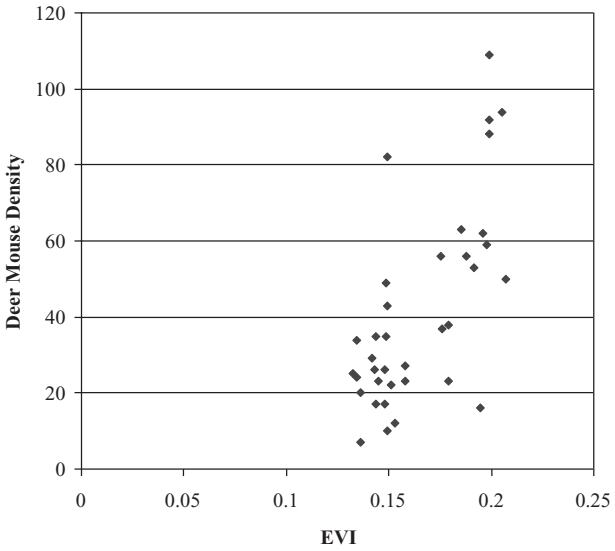

Figure 2 Deer mouse densities versus the enhanced vegetation index (EVI) on 10 June at a 1.0-year lag. The r^2 is 0.46 and significant at a level of 0.001.

with r^2 ranging from 0.21 to 0.29. There were significant correlations between NDWI at lags of 0 and 0.3 years and deer mouse density, with r^2 of 0.35 and 0.23.

Average 9 May NDVI and mouse density in the autumn across 12 study sites were computed and plotted in Fig. 3. The average 9 May NDVI increased 19% in 2005 from 2004, while mouse density doubled in the autumn of 2006 over the autumn of 2005. When average NDVI varied little between 2002 and 2003, there was little change in mouse density from 2003 to 2004.

Total number of infected mice and vegetation indices

Correlations between the number of infected mice and the vegetation indices across the time lags (0, 0.3, 1.0, 1.3 years) were also examined. Table 2 lists the maximum r^2 and the dates of the maximum r^2 for the number of infected mice and the vegetation indices. The results were very similar to those of mouse density. This may indicate that density and the number of infected mice are correlated.

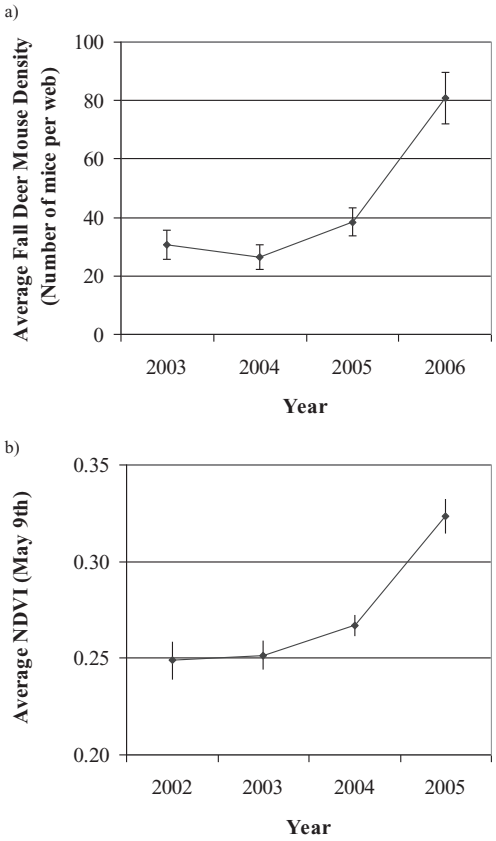

Figure 3 (a) Average autumn deer mouse density (number of mice per web ± SE) from 2003 to 2006. (b) Average normalized difference vegetation index (NDVI) on 9 May for 2002–05.

All vegetation indices had significant ($P < 0.01$) correlations with the number of infected mice at various time lags. For correlations between the number of infected deer mice and NDVI/EVI over various time lags, the 1.0- and 1.3-year lags yielded higher r^2 than the 0- and 0.3-year lags. Correlations were highest for EVI for a 1.0-year lag, with an r^2 of 0.53 and significance greater than 99.9%. Correlations were also high for NDVI for a 1.0- and 1.3-year lag, with r^2 of 0.36 and 0.43, respectively, and significance greater than 99.9%. There were significant correlations between NDWI at 0-, 0.3- and 1.0-year lags and the

Table 2 Maximum r^2 values and dates of maximum r^2 values between March and July for normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), visible atmospherically resistant index (VARI) and normalized difference water index (NDWI) regressed against the number of infected deer mice at 0-, 0.3-, 1.0- and 1.3-year lags.

Time lag (years)	Maximum NDVI r^2	Date of maximum NDVI r^2	Maximum EVI r^2	Date of maximum EVI r^2	Maximum VARI r^2	Date of maximum VARI r^2	Maximum NDWI r^2	Date of maximum NDWI r^2
0	0.18	6 March	0.11	10 June	0.23*	6 March	0.25*	6 March
0.3	0.09	6 March	0.03	10 April	0.12	6 March	0.16*	26 June
1	0.35*	10 June	0.53*	10 June	0.40*	10 June	0.20*	25 May
1.3	0.43*	10 June	0.36*	9 May	0.21*	9 May	0.04	9 May

*Significance ($P < 0.01$).

Table 3 Maximum r^2 values and dates of maximum r^2 values between March and July for normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), visible atmospherically resistant index (VARI) and normalized difference water index (NDWI) regressed against the prevalence of Sin Nombre virus (SNV) at 0-, 0.3-, 1.0- and 1.3-year lags.

Time lag (years)	Maximum NDVI r^2	Date of maximum NDVI r^2	Maximum EVI r^2	Date of maximum EVI r^2	Maximum VARI r^2	Date of maximum VARI r^2	Maximum NDWI r^2	Date of maximum NDWI r^2
0	0.12	7 April	0.05	9 May	0.28*	7 April	0.11	25 May
0.3	0.11	26 June	0.11	7 April	0.07	7 April	0.02	7 April
1	0.27*	26 June	0.25*	10 June	0.40*	7 April	0.19	7 April
1.3	0.06	28 July	0.03	28 July	0.05	7 April	0.03	6 March

*Significance ($P < 0.01$).

number of infected deer mice, with r^2 of 0.25, 0.16 and 0.20, respectively. Similar to the results for density, the maximum r^2 for NDWI and the number of infected mice was in March, while for NDVI and EVI the high significant maximum r^2 were in May and June.

SNV prevalence and vegetation indices

Correlations between SNV prevalence and the vegetation indices across the time lags (0, 0.3, 1.0, 1.3 years) were examined. Table 3 lists the maximum r^2 and the dates of the maximum r^2 of SNV prevalence and the vegetation indices.

Most of the correlations of prevalence and vegetation indices were not significant. The highest significant r^2 (0.40) was of VARI in June for a 1.0=year lag. The next highest r^2 was for NDVI with an r^2 of 0.28 with a 0-year lag. The next highest r^2 were of NDVI and EVI in June with a 1.0-year lag, with values of 0.27 and 0.25, respectively. NDWI did not have significant correlations with SNV prevalence for any time lag.

DISCUSSION AND CONCLUSIONS

This study examined the relationships between vegetation indices (NDVI, EVI, VARI and NDWI), deer mouse density, SNV prevalence and the number of infected deer mice (potential risk to humans) at various time lags with the goal of identifying useful vegetation indices and time lags for predicting the risk of hantavirus outbreaks. The results revealed reasonably good correlations between vegetation indices and deer mouse density and the number of infected deer mice. In contrast, correlations between vegetation indices and SNV prevalence were much weaker. Prevalence is calculated as the ratio of the number of infected mice and the overall number of mice, but may not necessarily be correlated with either the number of infected mice or the overall number of mice. For example, the prevalence on a study site with a low number of infected mice and low overall number of mice may be the same as the prevalence on a study site with a high number of infected mice and high overall number of mice.

Among four tested vegetation indices, NDVI and EVI yielded higher significant r^2 while NDWI had lower significant values. VARI produced the lowest r^2. The 1.0- and 1.3-year lags for NDVI and EVI provided the best correlations overall, which suggests that spring NDVI and EVI are better correlated with the following year's deer mouse abundance and the number of infected deer mice than the current year's. This finding corresponds to Glass *et al.'s* (2002) study, which found that there was an apparent 1-year lag between the 1997–98 El Niño and the increase in SNV prevalence in high-risk areas in the subsequent year. This was also the case in the 1993 hantavirus pulmonary syndrome outbreak in the south-western USA, which indicated that the previous 1991–92 El Niño event was the major climatic

Table 4 The diseases directly transmitted by rodents and their locations in the world.

Disease	Where the disease occurs
Hantavirus pulmonary syndrome	Throughout most of North and South America
Haemorrhagic fever with renal syndrome	Primarily in eastern Asia, Russia, Korea, Scandinavia, western Europe and the Balkans.
Lassa fever	West Africa
Leptospirosis	World-wide
Lymphocytic chorio-meningitis (LCM)	World-wide
Omsk haemorrhagic fever	Western Siberia
Plague	Western USA, South America, Africa, Asia
Rat-bite fever	World-wide: *Streptobacillus moniliformis* in North America and Europe, *Spirillum minus* in Asia and Africa
Salmonellosis	World-wide
South American arenaviruses (Argentine haemorrhagic fever, Bolivian haemorrhagic fever, Sabiá-associated haemorrhagic fever, Venezuelan haemorrhagic fever)	South America: parts of Argentina, Bolivia, Venezuela and Brazil
Tularemia	World-wide

factor leading to the outbreak. The dramatically increased rainfall is believed to result in greater food resources that permitted increase in the size of deer mouse populations and subsequently the human risk of hantavirus pulmonary syndrome (Parmenter *et al.*, 1993; Engelthaler *et al.*, 1999). This time lag is also consistent with the findings of the trophic cascade hypothesis (Yates *et al.*, 2002). The 1-year time lag provides great opportunities for applying satellite imagery to predict upcoming SNV outbreaks allowing preventive strategies to be deployed in a timely fashion.

While NDVI and EVI had the most significant r^2 at 1.0- and 1.3-year lags, NDWI had most significant r^2 at 0- and 0.3-year lags. The significant correlation between NDWI and deer mouse density occurred on 6 March. The most significant correlation between NDWI and the number of infected mice was also on 6 March. This indicates that early spring vegetation moisture may be a predictor of deer mouse density and the number of infected mice in the same spring and following autumn. Combinations of different indices and lags could be valuable for identifying the time windows for data collection for practical uses in monitoring rodent abundance and subsequent disease risk to humans.

Although there have been a few studies using remotely sensed data to predict hantavirus risk (Boone *et al.*, 2000; Glass *et al.*, 2000) most of them used single-date TM imagery. However, single-date image data make it difficult to capture the vegetation dynamics and their effect on disease dynamics. The high-temporal-resolution compositing satellite data eliminate the cloud cover and bad imagery, providing more reliable data. Glass *et al.* (2007) explored the use of January 2002 to April 2004 MODIS NDVI data to compare the seasonal patterns of vegetation growth in low- and high-risk hantavirus areas. The study showed that vegetation growth at high-risk sites started earlier and lasted longer than that at comparable lower risk sites, regardless of land cover. Marston *et al.* (2007) showed the advantages and associated potential of using time-series MODIS NDVI datasets to model rodent distributions over single-time NDVI datasets. The use of high-temporal-resolution MODIS imagery in epidemiological studies has been limited. This study demonstrated that the use of MODIS data in estimating rodent abundance and predicting disease risk to humans has potential. While our study area is in Utah, USA, MODIS data are freely available with world-wide coverage. This method is suitable for extension to other rodent-borne diseases around the globe. A few examples are leptospirosis and Argentine haemorrhagic fever. Herbreteau *et al.* (2006) explored the use of land-use maps and vegetation indices derived from advanced spaceborne thermal emission and reflection radiometer satellite imagery in studying leptospirosis dynamics in rice fields in Thailand. Porcasi *et al.* (2005) incorporated the advanced very-high-resolution radiometer NDVI as an environmental variable in a simple numerical model of rodent population dynamics and viral infection for the Junin virus. These studies have shown the usefulness of satellite-derived vegetation indices in predicting rodent-borne diseases. MODIS data may improve predictive models for many of the 11 types of rodent-borne diseases listed in Table 4 (CDC, 2009).

A predictive model utilizing NDVI or EVI and NDWI could be built to predict deer mouse abundance and hantavirus risk to humans. Other factors should be built into the model, such as temperature and the length of vegetation greenness. Tersago *et al.* (2009) found high summer and autumn temperatures in the preceding 1 and 2 years were related to high incidences of hantavirus disease in Belgium. Clement *et al.* (2009) reported that an increase in bank vole populations may be because a mild autumn in the previous year provides more food resources and therefore increases the survival rate in the winter. Longer vegetation greenness has been associated with areas of high HPS risk. Glass *et al.* (2007) found the NDVI in areas with high HPS risk had an early onset, with significantly higher levels of green vegetation that lasted longer than at lower-risk sites. In a large areal extent, other variables have been found useful in relating environmental conditions and rodent borne diseases, such as land use (Goodin *et al.*, 2006), vegetation types (Boone *et al.*, 2000) and elevation (Jay *et al.*, 1997).

While the results are promising, further study with more years of data would help to prove the long-term efficacy. Long-term studies are crucial for understanding the temporal patterns of virus host populations and identifying characteristics of reservoir ecology associated with outbreaks of human disease. In spite of their importance and utility, long-term studies of reservoir populations of zoonotic diseases are rare. They require continuous funding for many years, they are labour intensive, expensive and may not produce significant results in the short term (Mills *et al.*, 1999a,b). Our study collected a large number of field mouse trapping data that have not been seen in past hantavirus studies. During six field seasons from 2004 to 2006, over 4000 mice at 12 study sites were captured, recorded and released. These valuable field data provide great opportunities to understand mouse population ecology and hantavirus dynamics.

ACKNOWLEDGEMENTS

This research was supported by a NSF-NIH grant (EF 0326999). We thank J. Pearce, C. Clay, E. Lehmer, B. Wood and numerous crew members for the field data collection.

REFERENCES

Abbott, K.D., Ksiazek, T.G. & Mills, J.N. (1999) Long-term hantavirus persistence in rodent populations in Central Arizona. *Emerging Infectious Diseases*, **5**, 102–112.

Anderson, D.R., Burnham, K.P., White, G.C. & Otis, D.L. (1983) Density estimation of small-mammal populations using a trapping web and distance sampling methods. *Ecology*, **64**, 674–680.

Beck, P.S.A., Atzberger, C., Hogda, K.A., Johansen, B. & Skidmore, A.K. (2006) Improved monitoring of vegetation dynamics at very high latitudes: a new method for using MODIS NDVI. *Remote Sensing of Environment*, **100**, 321–334.

Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O., Schelde, K. & Thomsen, A. (2002) Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sensing of Environment*, **81**, 179– 193.

Boone, J.D., McGwire, K.C., Otteson, E.W., DeBaca, R.S., Kuhn, E.A., Villard, P., Brussard, P.F. & St Jeor, S.C. (2000) Remote sensing and geographic information systems: charting Sin Nombre virus infections in deer mice. *Emerging Infectious Diseases*, **6**, 248–258.

Buceta, J., Escudero, C., Rubia, F.J. & Lindenberg, K. (2004) Outbreaks of hantavirus induced by seasonality. *Physical Review E*, **69**, 021906.

Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1993) *Distance sampling: estimating abundance of biological populations*. Chapman and Hall, London.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford.

Burnham, K.P., Anderson, D.R. & Laake, J.L. (1980) Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, **72**, 1–202.

Calisher, C.H., Sweeney, W., Mills, J.N. & Beaty, B.J. (1999) Natural history of Sin Nombre virus in western Colorado. *Emerging Infectious Diseases*, **5**, 126–134.

Cantoni, G., Padula, P., Calderón, G., Mills, J.N., Herrero, E., Sandoval, P., Martínez, V., Pini, N. & Larrieu, E. (2001) Seasonal variations in prevalence of antibody to hantaviruses in rodents from southern Argentina. *Tropical Medicine and International Health*, **6**, 811–816.

Centers for Disease Control and Prevention (CDC) (1993) Outbreak of acute illness – Southwestern United States, 1993. *Morbidity and Mortality Weekly Report*, **42**, 421–424.

Centers for Disease Control and Prevention (CDC) (2009) *Diseases directly transmitted by rodents*. Available at: http://www.cdc.gov/rodents/diseases/direct.html (accessed October 2009).

Childs, J.E., Ksiazek, T.G., Spiropoulou, C.F., Krebs, J.W., Morzunov, S., Maupin, G.O., Gage, K.L., Rollin, P.E., Sarisky, J., Enscore, R.E., Frey, J.K., Peters, C.J. & Nichol, S.T. (1994) Serologic and genetic identification of *Peromyscus maniculatus* as the primary rodent reservoir for a new hantavirus in the southwestern United States. *Journal of Infectious Disease*, **6**, 1271–1280.

Childs, J.E., Krebs, J.W., Ksiazek, T.G., Maupin, G.O., Gage, K.L., Rollin, P.E., Zeitz, P.S., Sarisky, J., Enscore, R.E., Butler, J.C., Cheek, J.E., Glass, G.E. & Peters, C.J. (1995) A household-based, case-control study of environmental factors associated with hantavirus pulmonary syndrome in the southwestern United States. *American Journal of Tropical Medicine and Hygiene*, **52**, 393–397.

Clement, J., Vercauteren, J., Verstraeten, W.W., Ducoffre, G., Barrios, J.M., Vandamme, M., Maes, P. & Van Ranst, M. (2009) Relating increasing hantavirus incidence to the changing climate: the mast connection. *International Journal of Health Geographics*, **8**(1), doi:10.1186/1476-072X-8-1.

Custer, D.M., Thompson, E., Schmaljohn, C.S., Ksiazek, T.G. & Hooper, J.W. (2003) Active and passive vaccination against hantavirus pulmonary syndrome with Andes virus M genome segment-based DNA vaccine. *Journal of Virology*, **77**, 9894–9905.

Dearing, M.D., Previtali, M.A., Jones, J.D., Ely, P.W. & Wood, B.A. (2009) Seasonal variation in Sin Nombre virus infections in deer mice: preliminary results. *Journal of Wildlife Diseases*, **45**, 430–436.

Dennison, P.E., Roberts, D.A. & Peterson, S.H. (2007) Spectral shape-based temporal compositing algorithms for MODIS surface reflectance data. *Remote Sensing of Environment*, **109**, 510–522.

Engelthaler, D.M., Mosley, D.G., Cheek, J.E., Levy, C.E., Komatsu, K.K., Ettestad, P., Davis, T., Tanda, D.T., Miller, L., Frampton, J.W., Porter, R. & Bryan, R.T. (1999) Climatic and environmental patterns associated with hantavirus pulmonary syndrome, Four Corners region, United States. *Emerging Infectious Diseases*, **5**, 87–94.

Feldman, H., Sanchez, A., Murozonov, S., Spiropoulou, C.F., Rollin, P.E., Ksiazek, T.G., Peters, C.J. & Nichol, S.T. (1993) Utilization of autopsy RNA for the synthesis of the nucleocapsid antigen of a newly recognized virus associated with hantavirus pulmonary syndrome. *Virus Research*, **30**, 351–367.

Gao, B.C. (1996) NDWI – a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, **58**, 257–266.

Gitelson, A.A., Kaufman, Y.J., Stark, R. & Rundquist, D. (2002) Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, **80**, 76–87.

Glass, G.E., Cheek, J.E., Patz, J.A., Shields, T.M., Doyle, T.J., Thoroughman, D.A., Hunt, D.K., Enscore, R.E., Gage, K.L., Irland, C., Peters, C.J. & Bryan, R. (2000) Using remotely sensed data to identify areas of risk for hantavirus pulmonary syndrome. *Emerging Infectious Diseases*, **63**, 238–247.

Glass, G.E., Yates, T.L., Fine, J.B., Shields, T.M., Kendall, J.B., Hope, A.G., Parmenter, C.A., Peters, C.J., Ksiazek, T.G., Li, C.S., Patz, J.A. & Mills, J.N. (2002) Satellite imagery characterizes local animal reservoir populations of Sin Nombre virus in southwestern United States. *Proceedings of the National Academy of Sciences USA*, **99**, 16817–16822.

Glass, G.E., Shields, T.M., Cai, B., Yates, T.L. & Parmenter, R. (2007) Persistently highest risk areas for hantavirus pulmonary syndrome: potential sites for refugia. *Ecological Applications*, **17**, 129–139.

Goodin, D.G., Koch, D.E., Owen, R.D., Chu, Y., Hutchinson, J.S. & Jonsson, C.B. (2006) Land cover associated with hantavirus presence in Paraguay. *Global Ecology and Biogeography*, **15**, 519–527.

Herbreteau, V., Demoraes, F., Khaungaew, W., Hugot, J.P., Gonzalez, J.P., Kittayapong, P. & Souris, M. (2006) Use of geographic information system and remote sensing for assessing environment influence on leptospirosis incidence, Phrae province, Thailand. *International Journal of Geoinformatics*, **2**, 43–49.

Hjelle, B., Tórrez-Martínez, N., Koster, F.T., Jay, M., Ascher, M.S., Brown, T., Reynolds, P., Ettestad, P., Voorhees, R.E., Sarisky, J., Enscore, R.E., Sands, L., Mosley, D.G., Kioski, C., Bryan, R.T. & Sewell, C.M. (1996) Epidemiologic linkage of rodent and human hantavirus genomic sequences in case investigations of hantavirus pulmonary syndrome. *Journal of Infectious Disease*, **173**, 781–786.

Huete, A.R., Liu, H.Q., Batchily, K. & YanLeeuwen, W. (1997) A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, **59**, 440–451.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. & Ferreira, L.G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, **83**, 195–213.

Jay, M., Ascher, M.S., Chomel, B.B., Madon, M., Sesline, D., Enge, B.A., Hjelle, B., Ksiazek, T.G., Rollin, P.E., Kass, P.H. & Reilly, K. (1997) Seroepidemiologic studies of hantavirus infection among wild rodents in California. *Emerging Infectious Diseases*, **3**, 183–190.

Jin, S. & Sader, S.A. (2005) MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. *Remote Sensing of Environment*, **99**, 462–470.

Justice, C.O., Vermote, E., Townshend, J.R.G. *et al.* (1998) The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, 1228–1249.

Kuenzi, A.J., Morrison, M.L., Swann, D.E., Hardy, P.C. & Downard, G.T. (1999) A longitudinal study of Sin Nombre virus prevalence in rodents, southeastern Arizona. *Emerging Infectious Diseases*, **5**, 113–117.

Lehmer, E.M., Clay, C.A., Pearce-Duvet, J., St. Jeor, S. & Dearing, M.D. (2008) Differential regulation of pathogens: the role of habitat disturbance in predicting prevalence of Sin Nombre. *Oecologia*, **155**, 429–439.

Linthicum, K.J., Anyamba, A., Tucker, C.J., Kelley, P.W., Myers, M.F. & Peters, C.J. (1999) Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science*, **285**, 397–400.

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G. & Worthy, L.D. (2006) Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, **105**, 142–154.

Marston, C.G., Armitage, R.P., Danson, F.M., Giraudoux, P., Ramirez, A. & Craig, P.S. (2007) Spatial-temporal modeling of small mammal distributions using MODIS NDVI time-series data. *10th International Symposium on Physical Measurements and Spectral Signatures in Remote Sensing* (ed. by M.E. Schaepman, S. Liang, N.E. Groot and M. Kneubühler), *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **Vol. XXXVI**, Part 7/P9, 564. ISPRS, Davos, Switzerland.

Mills, J.N., Terry, L.Y. & Childs, J.E. (1995) Guidelines for working with rodents potentially infected with hantavirus. *Journal of Mammalogy*, **76**, 716–722.

Mills, J.N., Ksiazek, T.G., Peters, C.J. & Childs, J.E. (1999a) Long-term studies of hantavirus reservoir populations in the southwestern United States: a synthesis. *Emerging Infectious Disesases*, **5**, 135–142.

Mills, J.N., Yates, T.L., Ksiazek, T.G., Peters, C.J. & Childs, J.E. (1999b) Long-term studies of hantavirus reservoir populations in the southwestern United States, rationale, potential, and relevance for human health. *Emerging Infectious Diseases*, **5**, 95–101.

Nichol, S.T., Spiropoulou, C.F., Morzunov, S., Rollin, P.E., Ksiazek, T.G., Feldmann, H., Sanchez, A., Childs, J., Zaki, S. & Peters, C.J. (1993) Genetic identification of a novel hantavirus associated with an outbreak of acute respiratory illness in the southwestern United States. *Science*, **262**, 914–917.

Parmenter, R.R., Brunt, J.W., Moore, D.I. & Ernest, M.S. (1993) The hantavirus epidemic in the Southwest: rodent population dynamics and the implications for transmission of hantavirus-associated adult respiratory distress syndrome (HARDS) in the Four Corners region. *Sevilleta Publication*,

41, 1–45. Sevilleta Long Term Ecological Research (LTER) Project, University of New Mexico, Albuquerque, New Mexico, USA.

Parmenter, R.R., Pratap Yadav, E., Parmenter, C.A., Ettestad, P. & Gage, K.L. (1999) Incidence of plague associated with increased winter-spring precipitation in New Mexico. *American Journal of Tropical Medicine and Hygiene*, **61**, 814–821.

Pearce-Duvet, J.M.C., St Jeor, S.C., Boone, J.D. & Dearing, M.D. (2006) Changes in Sin Nombre virus antibody prevalence across seasons: the interaction between habitat, sex and infection in deer mice (*Peromyscus maniculatus*). *Journal of Wildlife Diseases*, **42**, 819–824.

Porcasi, X., Calderón, G., Lamfri, M., Gardenal, N., Polop, J., Sabattini, M. & Scavuzzo, C.M. (2005) The use of satellite data in modeling population dynamics and prevalence of infection in the rodent reservoir of Junin virus. *Ecological Modelling*, **185**, 437–449.

Qi, J. & Kerr, Y. (1997) On current compositing algorithms. *Remote Sensing Reviews*, **15**, 235–256.

Root, J.J., Calisher, C.H. & Beaty, B.J. (1999) Relationships of deer mouse movement, vegetative structure, and prevalence of infection with Sin Nombre virus. *Journal of WildLife Disease*, **35**, 311–318.

Rouse, J.W., Haas, R.H., Schell, J.A. & Deering, D.W. (1973) Monitoring vegetation systems in the Great Plains with ERTS. *Proceedings of the Third ERTS Symposium*, NASA Special Publication 351, Vol. 1. , pp. 309–317.

Tersago, K., Verhagen, R., Servais, A., Heyman, P., Ducoffre, G. & Leirs, H. (2009) Hantavirus disease (nephropathia epidemica) in Belgium: effects of tree seed production and climate. *Epidemiology and Infection*, **137**, 250–256.

Thomson, M.C., Connor, S.J., Milligan, P.J.M. & Flasse, S.P. (1997) Mapping malaria risk in Africa: what can satellite data contribute? *Parasitology Today*, **13**, 313–318.

Thomson, M.C., Connor, S.J., D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M. & Greenwood, B. (1999) Predicting malaria infection in Gambian children from satellite data and bed net use surveys: the importance of spatial correlation in the interpretation of results. *American Journal of Tropical Medicine and Hygiene*, **61**, 2–8.

Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J.Y., Salas, W. & Moore, III, B. (2006) Mapping paddy rice agriculture in south and southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, **100**, 95–113.

Yates, T.L., Mills, J.N., Parmenter, C.A., Ksiazek, T.G., Parmenter, R.R., Castle, J.R.V., Calisher, C.H., Nichol, S.T., Abbott, K.D., Young, J.C., Morrison, M.L., Beaty, B.J., Dunnum, J.L., Baker, R.J., Salazar-Bravo, J. & Peters, C. (2002) The ecology and evolutionary history of an emergent disease: hantavirus pulmonary syndrome. *BioScience*, **52**, 989–998.

BIOSKETCHES

Lina Cao has obtained her PhD in Geography at the University of Utah. She is currently a postdoctoral scholar at University of California, Berkeley. Her research interests are spatial analysis, geographic information systems and remote sensing in public health.

Thomas J. Cova is Associate Professor of Geography and Director of the Center for Natural and Technological Hazards at the University of Utah. His research focuses on hazards and emergency management, transportation and location science and geographic information systems and science.

Philip E. Dennison is Associate Professor of Geography and Associate Director of the Center for Natural and Technological Hazards at the University of Utah. His research is centred on remote sensing of vegetation physiology and phenology, imaging spectroscopy, wildfire and fire danger modelling, vegetation and the carbon cycle, natural hazards and physical geography.

Editor: Thomas Gillespie