Int J Appl Earth Obs Geoinformation 64 (2018) 249–255

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Research Paper

Detecting plague-host abundance from space: Using a spectral vegetation index to identify occupancy of great gerbil burrows

Liesbeth I. Wilschut[a,b,*], Johan A.P. Heesterbeek[b], Mike Begon[c], Steven M. de Jong[a], Vladimir Ageyev[d], Anne Laudisoit[c,e], Elisabeth A. Addink[a]

[a] Utrecht University, Department of Physical Geography, Heidelberglaan 2, PO Box 80115, 3508 TC Utrecht, The Netherlands
[b] Utrecht University, Faculty of Veterinary Medicine, Yalelaan 7, 3584 CL Utrecht, The Netherlands
[c] Ecology Evolution and Genomics of Infectious Disease Research Group, Institute of Integrative Biology, The University of Liverpool, Liverpool, UK
[d] Anti-Plague Institute, M. Aikimbayev's Kazakh Science Centre for Quarantine and Zoonotic Diseases, 14 Kapalskaya Street, Almaty 050074, Kazakhstan
[e] Evolutionary Biology Group, University of Antwerp, 171, Groenenborgerlaan, 2020 Antwerpen, Belgium

ARTICLE INFO

Keywords:
Plague
Yersinia pestis
Infectious disease
Great gerbil
Population abundance
NDVI
Object-based image analysis
Segmentation
Random Forest

ABSTRACT

In Kazakhstan, plague outbreaks occur when its main host, the great gerbil, exceeds an abundance threshold. These live in family groups in burrows, which can be mapped using remote sensing. Occupancy (percentage of burrows occupied) is a good proxy for abundance and hence the possibility of an outbreak. Here we use time series of satellite images to estimate occupancy remotely.

In April and September 2013, 872 burrows were identified in the field as either occupied or empty. For satellite images acquired between April and August, 'burrow objects' were identified and matched to the field burrows. The burrow objects were represented by 25 different polygon types, then classified (using a majority vote from 10 Random Forests) as occupied or empty, using Normalized Difference Vegetation Indices (NDVI) calculated for all images. Throughout the season NDVI values were higher for empty than for occupied burrows.

Occupancy status of individual burrows that were continuously occupied or empty, was classified with producer's and user's accuracy values of 63 and 64% for the optimum polygon. Occupancy level was predicted very well and differed 2% from the observed occupancy. This establishes firmly the principle that occupancy can be estimated using satellite images with the potential to predict plague outbreaks over extensive areas with much greater ease and accuracy than previously.

1. Introduction

Plague (*Yersinia pestis* infection) is a flea-borne zoonotic disease that is infamous for inducing three pandemics in the past two millennia (Gage & Kosoy, 2005). Currently, the plague agent circulates in rodent populations mainly in Africa, the Americas and Asia, and causes human deaths predominantly in Africa (World Health Organization, 2004). In Kazakhstan, plague occurs, and has been extensively studied, in populations of its main local host, the great gerbil (*Rhombomys opimus*). Great gerbils are burrowing, mainly folivorous rodents that live in semi-desert environments. They live in family groups usually consisting of one male, one or more females and their offspring, and they occupy one burrow per family (Randall et al., 2005).

Presence and absence data of plague from these populations show that plague requires a minimum abundance of great gerbils to be able to spread successfully – the so-called abundance threshold for plague (Davis et al., 2004). The number of gerbils in a burrow varies; a study from 2005, for example, found mean group sizes varied from 3.9 in 1996 to 13.4 in 1998 (Randall et al., 2005). Nonetheless, Davis et al. (2004) found that to monitor fluctuations in great gerbil abundance and predict plague outbreaks accurately, it is not necessary to know the exact number of gerbils in one burrow. Rather, the percentage of the burrows occupied (referred to as the occupancy level) is an effective proxy for abundance and also easier to measure. This is because great gerbils living in the same burrow tend to have the same disease status, i.e. plague is transmitted easily between family members (Davis et al., 2007). Indeed, these occupancy levels can be used to predict plague outbreaks two years in advance (Davis et al., 2004), and they define a 'percolation threshold', which, when exceeded, allows plague to spread between occupied burrows across the landscape (Davis et al., 2008; Reijniers et al., 2012).

In the past, and especially since the 1950s, field sampling has been carried out in Kazakhstan to monitor the occupancy level in the field as a means of estimating great gerbil abundance with a view to controlling

* Corresponding author at: Utrecht University, Department of Physical Geography, Heidelberglaan 2, PO Box 80115, Utrecht, 3508 TC, The Netherlands.
E-mail address: liesbethwilschut@xs4all.nl (L.I. Wilschut).

http://dx.doi.org/10.1016/j.jag.2017.09.013
Received 13 July 2017; Received in revised form 23 September 2017; Accepted 25 September 2017

human plague (Davis et al., 2004). This is expensive in terms of time and manpower, which in turn sets limits on the spatial extent of surveillance. Recently, remote sensing methods were developed that allow highly accurate identification of the presence of burrows (Addink et al., 2010; Wilschut et al., 2013a). However, occupied and unoccupied burrows could not be distinguished with these methods, and occupancy levels (abundance) therefore could not be computed.

The desirability of monitoring plague through tracking gerbil occupancy remotely is thus a focused example of the more general challenge of monitoring animal abundance from space. Although sparse, some examples do exist. Fretwell et al. (2017) estimated the size of an albatross population by counting them manually in 30cm-resolution Worldview-3 imagery. Fretwell et al. (2014) counted whales by supervised and unsupervised classification. And Laliberte and Ripple (2003) counted cattle in an IKONOS image by a computer-aided approach. However, no examples were found on the indirect mapping of a species by detecting traces of an animal like the burrows they built and to determine whether the animals are still present.

Here we take up this more general challenge with the aim ultimately, in this particular case, of making medically relevant predictions of plague risk. Thus, we use high-spatial-resolution remote-sensing images to estimate the gerbils' occupancy level, aiming to distinguish occupied from empty burrows using variables derived from the images, and subsequently to develop an algorithm to accurately classify a burrows' occupancy status. To detect a difference in burrow occupancy, we use the behavioural characteristics of the great gerbils. As great gerbils are herbivores, they remove and eat plants and their roots from the surface of their burrow and from its surroundings. This removal of vegetation over the season may be visible in high-spatial-resolution imagery, by using vegetation indices such as the Normalized Difference Vegetation Index (NDVI). As there is a natural change in the NDVI curve of vegetation between spring and autumn, we hypothesize that this curve is different for occupied and empty burrows.

Two objectives are addressed:

1 The first is to assess the NDVI trend of burrow occupancy classes from April to August
2 The second is to use the information obtained to develop an algorithm, using Random Forests, to classify a burrows' occupancy status, allowing occupancy levels to be estimated from satellite imagery.

2. Methods

2.1. Field data

The study area is located in eastern Kazakhstan, south of Lake Balkhash. Data collection was carried out in 38 squares of 250 by 250 m and six squares of 200 by 200 m in April 2013 and repeated in September 2013. All squares were located in one so-called 'sector' (Fig. 1), a monitoring unit of the plague monitoring stations in Kazakhstan, approximately 9.3 km by 9.8 km in size (Wilschut et al., 2013b, 2015). The squares are all located in the older part of the river Ili delta consisting of floodplains covered by river sediments, scattered sand dunes and semi-arid vegetation. The diversity of the floodplain was reflected in the square locations. The inaccessibility of the terrain required all squares to be in the vicinity (< 1 km) of sand tracks. All burrows inside the squares were mapped by field observation, deriving coordinates from a GPS at the "ecological centre" of the burrow – the location of most intensive visible great gerbil activity. Burrow diameters were measured in two perpendicular axes crossing at the ecological centre (Wilschut et al., 2013a). The occupancy status of each burrow was determined using a standard protocol described in Wilschut et al. (2015) from signs of foraging activity, scent marks, presence of fresh faeces and the recent clearance of burrow entrances (indicated by the presence of freshly turned-over soil). Burrows were classified as either occupied or empty, and all burrows were given a unique code. In addition, the diameter of each burrow was measured.

2.2. Satellite data and pre-processing

Five 1.8 m resolution ortho-rectified Worldview-2 images (DigitalGlobe, 2014) with spectral bands Blue, Green, Red and NIR were acquired on April 12th, May 6th, June 5th, July 10th and August 15th 2013. All images were converted to top-of-atmosphere reflectance (Chander et al., 2009). The five images were geo-rectified and subsequently resampled to match the April image using automatic image registration. During evaluation of the data, the July image appeared unsuitable for usage because of many clouds, shade and haze in the image. It was therefore excluded from the analysis. In order to evaluate the overall NDVI trend, the mean NDVI value of the entire sector was calculated for each of the four months.

2.3. Evaluation of NDVI values of burrow-occupancy classes

During the two field campaigns, 872 burrows were mapped and categorized in the field as either occupied or empty in April 2013 and in September 2013. This allowed four categories of change between the two months: occupied–occupied (*oo*) in both April and September; empty–empty (*ee*); empty in April-occupied in September (*eo*) and occupied in April-empty in September (*oe*).

In order to assess the NDVI trend of burrow classes, first the area occupied by a burrow was identified by buffering the field coordinates of each burrow centre with the average radius of all burrows. The average burrow diameter mapped in the field in September was 17.8 m. Hence, a circular buffer of 9 m radius was used. The NDVI was calculated for all burrows in all images (April–August 2013). Then, the mean NDVI values for all the burrow-circles were calculated for each month.

2.4. Overall burrow identification using remote sensing

To identify the great-gerbil burrows (both occupied and empty) in the sector in the satellite images, the semi-automatic classification method as described in a previous study was used (Wilschut et al., 2013a), with some minor adaptations. Here, this method is only described briefly. First, the field data set was divided into two parts: 50% was used for training the classifier and 50% was used for validation of the burrow classification results. Then, the April 2013 image was segmented into objects using multi-resolution segmentation (Trimble, 2011) at 15 different spatial scales to be able to identify the spatial scale with the highest classification accuracy after classification. Next, the Random Forest classifier (Breiman, 2001) was used to classify the objects in the image as burrow or non-burrow objects. To validate the classification, several indicators were used. First, producer's, user's and overall accuracies were calculated (Lillesand et al., 2004). Producer's accuracy is the percentage of validation burrows that are correctly classified. The user's accuracy is the percentage of all validation objects classified as burrows that are actually burrows. The overall accuracy is the percentage of classified validation objects (burrows and non-burrow areas) that are classified correctly. Based on the field data, the densities of burrows in the squares were calculated. Then, the overall ratio between the density predicted by the classification and the density observed in the field was calculated to determine the accuracy of the density of the created burrow maps. Finally, the number of objects that intersected burrows multiple times was calculated to determine whether the size of the image objects resembled the size of the field burrows. The optimal segmentation scale was selected by choosing the scale where the combination of the validation variables reached its optimum. The objects that were correctly classified as burrows were processed further.