

An early warning system for Crimean-Congo haemorrhagic fever seasonality in Turkey based on remote sensing technology

Agustin Estrada-Peña¹, Zati Vatansever², Aysen Gargili³, Turan Buzgan⁴

¹*Department of Parasitology, Faculty of Veterinary Medicine, University of Zaragoza, Zaragoza, Spain;*

²*Department of Parasitology, Faculty of Veterinary Medicine, University of Ankara, Ankara, Turkey;*

³*Cerrahpasa Medical Faculty, Istanbul, Turkey;* ⁴*Primary Health Care Division, Ministry of Health, Ankara, Turkey*

Abstract. In the last few years, Crimean-Congo haemorrhagic fever (CCHF) has been reported as an emerging tick-borne disease in Turkey. This paper deals with the preparation of an early warning system, aimed to predict the beginning of the CCHF season in Turkey based on a clear, simple and repeatable remotely-sensed signal. Decadal (mean of 10 days) values of the normalized difference vegetation index (NDVI) at 1 km resolution were used on a set of 952 confirmed, accurately geo-referenced, clinical cases between 2003 and 2006. A prerequisite is that the signal should be observable between 2 and 3 decadals before a given moment of the season to be of value as early warning. Decadals marking the 10th percentile or the 25th quartile in the frequency distribution of case reporting were selected as markers for the beginning of season of risk. Neither raw nor accumulated decadal NDVI signals were able to predict the onset of this season. However, when we defined the NDVI anomaly (NDVIa) as the positive difference between decadal NDVI values and the average for the previous year, this standardized measure gave a homogeneous overview of the changes in the NDVI signal producing a NDVIa slope for the decadals 10 to 13 that was always greater than 0. We conclude that observing this slope over time can be used as an early-warning system as it would predict the build-up of the number of cases 20 days in advance with an accuracy of 82% (10th percentile) or 98% (25th quartile).

Keywords: Crimean-Congo haemorrhagic fever, early warning, Turkey, normalized difference vegetation index, remote sensing.

Introduction

The early identification of an epidemic infectious disease is an important first step towards implementing effective interventions to control the disease and reduce the resulting impact on humans. Usually, however, epidemics are well under way before the authorities are notified and control meas-

ures are put in place. Because the geographical and seasonal distributions of many infectious diseases are inherently linked to the climate, the possibility of using climate parameters as predictive indicators in early-warning systems (EWS) has long been a focus of interest. During the 1990s, a number of factors led to increased activity in this field, including significant advances in data availability, epidemiological modelling, information technology, and the implementation of successful EWS outside the health sector.

The potential of applying remote sensing (RS) to epidemiological studies has long been recognized. The principal goal of this approach is to map the distribution of a disease, often by mapping the distribution of its vector, as the case may be. The pre-

Corresponding author:

Agustin Estrada-Peña

Department of Parasitology

Faculty of Veterinary Medicine

University of Zaragoza

Miguel Servet 177, 50013 Zaragoza, Spain

Tel. +34 976 7615 558; Fax +34 976 761 612

E-mail: aestrada@unizar.es

requisite for all epidemiological applications of RS is a close association between the life history of a parasite or vector species and environmental features that can be observed remotely (Hugh-Jones, 1989). Previous reviews of RS focused on the problems of mosquito and malaria control, and the study of the range of arthropod-borne diseases have explored the relationships between remotely-sensed surrogates for the ecological determinants of vector distribution and disease abundance (Hay et al., 1998). However, less progress has been made in the prediction of the seasonality of the disease using remotely-sensed features. Thomson et al. (1996, 1997) studied the potential of NOAA-AVHRR and Meteosat-High Resolution Radiometer data to predict the malaria risk in The Gambia. Although clear relationships were demonstrated between satellite sensor data and meteorological variables associated with malaria transmission, it was problematic to predict how these would affect the number of adult mosquitoes owing to the complexity of the interrelationships. This situation is further complicated by sociological factors. Despite these difficulties, predictions of malaria seasonality (the combination of disease risk in space and time) in Kenya have been achieved by establishing relationships between cases of childhood malaria and contemporary satellite imagery (Hay et al., 1998).

Recent ecological studies have highlighted the relevance of the normalized difference vegetation index (NDVI) as a value describing vegetation performance and its condition on a real-time basis. The NDVI is derived from the red and near-infrared reflectance ratio reflected by the vegetation and captured by the sensor of the satellite. The relationship between NDVI and vegetation productivity is well established, and the link between this index and the fraction of absorbed photosynthetic active radiation intercepted has been well documented (i.e. Sellers et al., 1992).

Crimean-Congo haemorrhagic fever (CCHF) is a viral infection that occurs in parts of Africa, Asia, the Middle East and south-eastern Europe (Hoogstraal, 1979). The virus belongs to the genus

Nairovirus of the Bunyaviridae family and causes severe disease in humans, with a fatality rate ranging from 3 to 30% (Whitehouse, 2004). Its wide geographic distribution largely coincides with the distribution of the *Hyalomma* vector ticks (Hoogstraal, 1979). Human beings become infected either through the bite of infected *Hyalomma* ticks, or by contact with a patient with acute phase CCHF, or by contact with blood or tissues from viraemic livestock.

After decades of serologic evidence, Turkey is experiencing the emergence of a new situation. Outbreaks of the disease were first reported in 2002 from the northern mountainous area and eastern Anatolia (Ergonul et al., 2004; Karti et al., 2004; Bakir et al., 2005; Ozkurt et al., 2006). By the end of 2006, there were as many as 1,103 CCHF confirmed cases recorded by the Ministry of Health of Turkey. In a previous paper (Estrada-Peña et al., in press), we have demonstrated that the risk areas for CCHF in Turkey are associated with sites climatically highly suitable for the main tick vector, *Hy. m. marginatum*, and with large habitat fragmentation. The aim of this paper is to further validate the usefulness of NDVI values and to provide, as an additional preventive tool, an EWS for the prediction of the beginning of the CCHF season in Turkey.

Material and methods

Study area

Turkey extends from 36° to 42° North and from 26° to 45° East covering an area of 783,562 km². The country is located in the Mediterranean macroclimate zone between the temperate and sub-tropical zones in the west of the Asian continent. This region is characterized by widely diverse regional and/or seasonal variations, ranging from extremely harsh conditions to very hot, dry summers. The southern and western parts of the country lie under the influence of the Mediterranean climate with hot

and dry summers and cool and rainy winters. The climatic conditions on the Black Sea coast are colder and rainier. Northeast Anatolia has the characteristics of a continental climate with long and very cold winters and short, relatively cold summers. The central Anatolian plateau is under the influence of a steppe climate with arid and hot summers and cold winters.

Case notifications

The CCHF cases reported to the Turkish Ministry of Health and entered into their database were used as the basis for this study. This database includes detailed information at the village level for each case diagnosed in Turkey between 2002 and 2006. However, because the low number of cases in 2002 ($n = 17$) it was decided to use only cases recorded between 2003 and 2006. Each case was confirmed by the Refik Saydam National Hygiene Center in Turkey using standard serological and molecular polymerase chain reaction (PCR)-based tests. A total of 952 cases were adequately geo-referenced at a resolution of 1 km. The date of hospitalization was used as an unambiguous marker of the onset of the disease. The reported time of a tick bite is inadequate as marker for the beginning of the infection since many infected people do not notice the tick bite or are bitten several times. The date of hospitalization for each case was ascribed to the corresponding decadal (10-day interval) of each year. A grid of polygons of 10 km diameter was generated to cover the entire country and cases were spatially allocated to the covering hexagon in question. The median of the decadal for hospital application of cases included into each hexagon was used as marker of seasonality in that particular portion of territory.

RS data

Vegetation data (i.e. NDVI) were used to find an adequate marker for the beginning of the season for CCHF cases. We used decadal intervals of

images taken over the region by the SPOT satellite at 1 km resolution from 2002 to 2006. Spurious and contaminated pixels were eliminated according to Hay et al. (2006) and the complete series was smoothed using Fourier series analysis. Further analyses were performed only on the smoothed data series. First, we used NDVI values to produce a classification of vegetation in terms of categories according to the rationale put forward by Pettorelli et al. (2005). This was done to identify any association of cases with definite vegetation classes which would help in further analysis. Averaged NDVI values for the 2002-2006 period were used to produce 36 decadal layers from which categories were further obtained. Classification was performed in ERDAS Imagine 8.7 using an unsupervised classification technique and the ISO-DATA algorithm (Odiit et al., 2005) to identify 50 spectral signatures. Those were analysed to identify spectrally distinct signatures using the transformed divergence measure. A signature was merged, deleted or retained depending on the number of other signatures it was close to. This process was performed until 15 distinct signatures remained, then used as input for a maximum likelihood supervised classification. A 3×3 fuzzy filter was further applied to remove small patches comprising few pixels, which were assigned to the most probable neighbouring category.

Four parameters were derived from NDVI, namely:

- (i) raw NDVI values of each decadal;
- (ii) the sum of NDVI values for each decadal since the beginning of each year;
- (iii) anomalies of the NDVI at a given decadal; and
- (iv) slope of the NDVI anomalies observed at four consecutive decadals.

The NDVI anomaly (NDVIa) is defined as the positive difference between NDVI values for each single pixel at a given decadal and the average of the previous year. The rationale of this calculation is derived from the different seasonal amplitude of each vegetation category, crude NDVI values having very different values at the same moment of the year for different categories. NDVIa is thus a standard-

ized measure of how much NDVI values deviate from the average of the previous year giving a homogeneous overview of the changes in NDVI signal.

Statistical analysis

The grid of hexagons was further subjected to statistical analysis (software Statistica 6.0). Other than decadal of case reporting, hexagons were loaded with the dominant vegetation category inside a given polygon, number of different categories inside each hexagon, raw and accumulated NDVI values at each decadal, as well as NDVIa, and slope of NDVIa at every four consecutive decadal. The purpose was to detect a clear, repeatable and simple signal marking the beginning of the CCHF cases season. A prerequisite was that the signal should be observable between 2 and 3 decadal before a given moment of the risk season to be of value as an early-warning marker. Decadal marking the 10th percentile or the 25th quartile in the frequency distribution of case reporting were selected as markers for the beginning of the season of risk.

A one-way analysis of variance (ANOVA) was used to check if clinical cases occurred in defined

vegetation categories. The null hypothesis was rejected at values of $P < 0.05$. Linear regressions were performed between raw or accumulated NDVI values, NDVIa, or slope of NDVIa and the decadal of case-reporting to check for a threshold that would mark the beginning of risk season at the marks stated before.

Results

CCHF cases were found to be unevenly distributed in the different arbitrary vegetation categories with high reporting of cases within defined categories (ANOVA, $P < 0.001$). Almost half of the cases (Table 1 and Fig. 1) accumulated into category 8, with an additional 39% in the categories 5, 6 and 12. These categories of vegetation cover a broad geographical range in the country and there are wide areas with the same vegetation features where no cases have been reported (Fig. 1). Seasonality of reporting is clearly unimodal (Fig. 2), most cases being concentrated between the decadal 14 and 21. The 10th percentile of case reporting is located around decadal 13 and 14, with the 25th quartile at decadal 15 and 16, for

Table 1. Distribution of the number of CCHF cases in Turkey in the period of 2003-2006 according to the vegetation categories derived from NDVI values. The mean and median values refer to the distribution of frequencies of case reporting in the decadal.

Vegetation category	Number of cases	%	Mean	10 th percentile	25 th quartile	Median
1	Water	-	-	-	-	-
2	2	0.2	22.0	17	17	22
3	22	2.3	16.8	12	14	18
4	50	5.2	17.7	14	15	18
5	193	20.3	17.6	14	16	18
6	99	10.4	17.2	13	15	18
7	3	0.3	13.7	11	11	12
8	450	47.3	17.2	13	15	17
9	16	1.7	18.1	14	17	19
10	21	2.2	17.6	13	16	18
11	0	0	-	-	-	-
12	78	8.2	17.6	13	16	18
13	5	0.5	15.4	8	13	17
14	5	0.5	20.0	17	17	20
15	8	0.8	16.5	11	17	17
Total	952	-	17.4	13	16	18

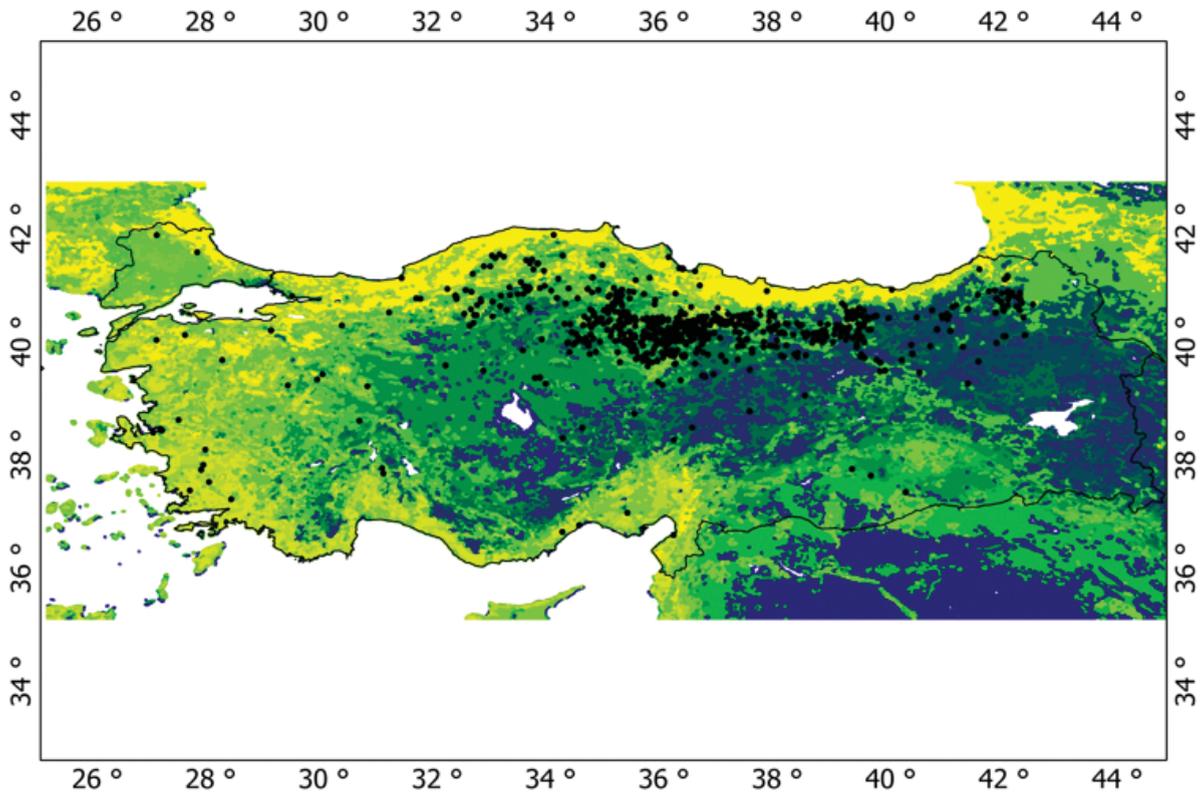


Fig. 1. The geographical distribution of 15 vegetation categories in Turkey, obtained from Fourier smoothed NDVI series of data (2003-2006) together with the distribution of CCHF cases over the same period.

the vegetation categories 5, 6 and 12 (where most cases are reported). Since it was decided to attempt identifying an early-warning system based on a clear NDVI signal at least 1 month in advance, the NDVI values for decadal 9 to 13 were thereafter explored.

Raw and accumulated NDVI values at the decadal 9 to 11 were uncorrelated with case seasonality (Figs. 3 and 4) and there was a broad variability in NDVI values within the same decadal of reporting and unrelated with the category of vegetation ($P = 0.324$). NDVI_a for the decadal 9 was close to 0 in all cases and there was also a broad variability for the decadal 10 to 13 (Fig. 5). However, there was a continuously increasing trend in NDVI_a for all locations so the slope of the NDVI_a for the decadal 10 to 13 was always greater than 0. The use of the slope of

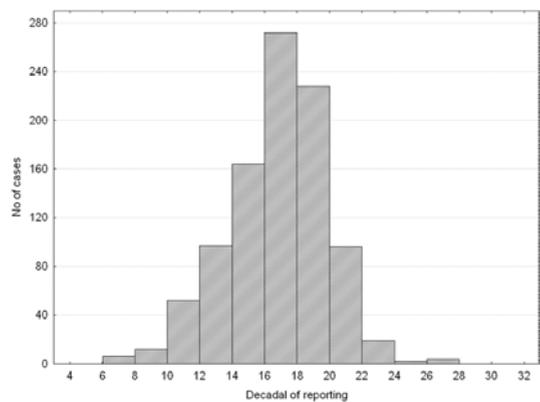


Fig. 2. Distribution of the number of CCHF cases in Turkey in the period of 2003-2006 according to the decadal of reporting. All data of the series were used to produce this averaged distribution.

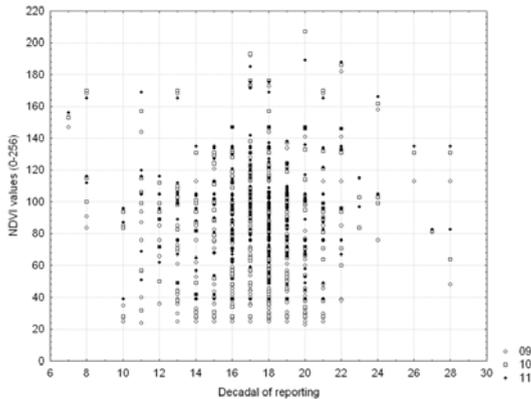


Fig. 3. Relationships between the raw NDVI values (0-256) in the decadals 9, 10 and 11 and the decadal of reporting of each case. The numbers refer to the decadal.

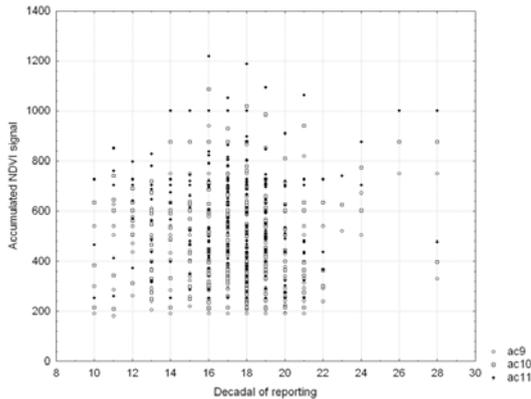


Fig. 4. Relationships between the accumulated NDVI values in the decadals 9, 10 and 11 (beginning on 1 January) and the decadal of reporting of each case. The numbers refer to the decadal.

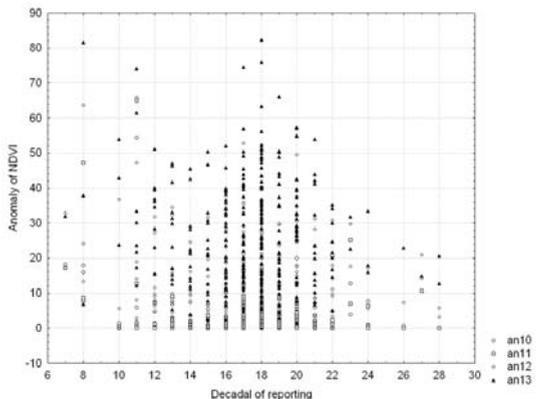


Fig. 5. Relationships between the anomaly in NDVI values in the decadals 10 to 13 with the decadal of case reporting. The numbers refer to the decadal.

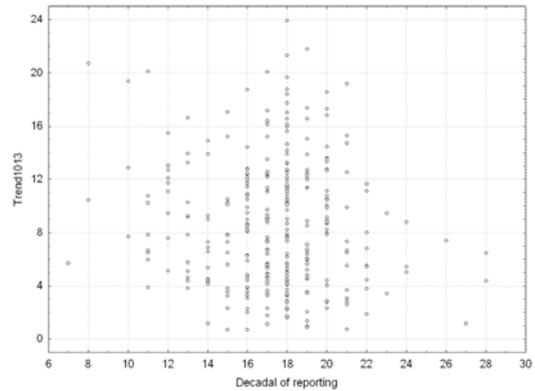


Fig. 6. Relationships between the slope of the anomaly on NDVI values in the four consecutive decadals 10 to 13 and the decadal of reporting of each case. Independently of the decadal of reporting, the slope of these NDVI values is always positive and associated with the admission activity in hospitals.

NDVIa for a specified period of time, as applied to the 10th percentile, would mark the onset of the risk season with an accuracy of 82.3%. The use of such a marker, as applied to the 25th quartile, would mark the beginning of the season with an accuracy of 97.9% (see Fig. 6). That means that warning of the risk can be issued 20 days in advance, using the slope of NDVIa in the decadals 10 to 13, as applied to the prediction of 75% of cases, with a high level of accuracy. These results were obtained for the whole series of data but they were also applied to the individual years in the period from 2003 to 2006 to check for repeats or inconsistencies. The lowest accuracy in the prediction of the onset of risk was obtained for the year 2003 (91%), and the highest in 2006 (98%).

Discussion

An important objective in using RS data in epidemiology is to map the distribution of a disease, clues to which can be found by mapping the distribution of its vectors (Hay, 1997). To predict where diseases might occur, it is necessary to understand patterns in geographical distribution and how these change over time (Rogers and Randolph, 2003). For many pathogens, especially the vector-borne ones,

transmission is seasonal. The purpose of this paper is to promote the use of remotely-sensed features of vegetation to issue an EWS for the prevention of the population at risk for CCHF in Turkey. We have already demonstrated (Estrada-Peña et al., in press) that the current geographical extent and the incidence of CCHF cases in Turkey are well correlated with the presence of suitable habitats for the main tick vector, *Hy. m. marginatum* (derived from monthly temperatures and NDVI values) together with a high degree of habitat patchiness (fragmentation), which increase the risk for contact between humans and the infected ticks. As temperatures and habitat fragmentation play an important role in the overall risk for the disease, the NDVI-derived vegetation categories alone are not useful for prediction, neither of the geographic range of CCHF in Turkey, nor of the incidence of the disease. Our study thus concentrated on the development of a method to detect in advance the onset of the season of risk and not to correlate a given NDVI value with the temporal incidence of the disease.

While the predictive mapping of some vector-disease systems need the inclusion of a combination of temperatures and NDVI values (e.g. schistosomiasis; see Kristensen et al., 2001; Malone et al., 2001) good correlates have been found between disease incidence rates and NDVI seasonal variation (e.g. malaria; see Hay et al., 1998, 2000). Furthermore, Rogers and Randolph (1991) have shown that the distribution and abundance of tsetse flies (the vector for sleeping sickness) can be related to NDVI. However, these authors point out that the abundance of tsetse flies alone are not the best indicator of risk because certain types of land cover may be of greater importance, due to human activity, and promote higher levels of human-vector contact.

Our aim has been to avoid complex statistical calculations based on land surface temperatures, ground moisture, atmospheric water vapour or rainfall and instead search for a clear and simple signal. Since vegetation dynamics and local climate are intrinsically linked, the approach is based on the rationale of using NDVI as a proxy for arthropod

seasonal activity. Many factors affect NDVI variations within a pixel: together with topography and altitude, plant architectural arrangement, interactions with canopy cover, height, composition of species, vegetation vigour, leaf properties and vegetation stress are some factors that can significantly affect the remotely-sensed information (Pettorelli et al., 2005). In our study, neither raw, nor accumulated NDVI values could be used to mark the onset of the CCHF season because the high variability observed in different vegetation categories at the periods of time investigated. However, the slope of NDVI_a (defined as the change in NDVI signal as compared to the average of the previous year) was consistently positively correlated with the beginning of CCHF transmission in Turkey with an accuracy ranging between 82% (for the 10th percentile) and 98% (for the 25th quartile). Some early cases in the year remain outside of the temporal window predicted with this method, probably because of local conditions that are below the resolution of this study.

The suite of ecological factors that influence CCHF transmission may vary across areas in Turkey and even in the whole European and central Asian regions where the disease occurs. Probably, the wild fauna is one of the most important culprits because of its role as virus reservoir (Hoogstraal, 1979). Pathogens, vectors, reservoirs and hosts are all embedded within ecological communities within which they interact, directly or indirectly, with other species. The number and interaction of dependent factors result in a formidable complexity and the importance of landscape composition (number and types of patches) and configuration (spatial relationships among patches) to disease dynamics is only beginning to be explored. Landscape structure has a strong potential to influence disease dynamics through its impact on both abiotic conditions and species interactions that are important for disease distribution and prevalence (Perkins et al., 2006). As diseases invade new areas or change distribution along with climate and land-use modifications, a major research challenge will be to determine the level of ecological complexity that is necessary to

predict spatial dynamics accurately (Rosà et al., 2003) and hence case incidence.

Considerably more research needs to be conducted to define the empirical relationships between ecology of the tick vector(s) and their hosts, contact with humans, habitat fragmentation, disease risk and remotely-sensed variables. More information about these relationships will undoubtedly provide baseline data to model tick population dynamics in real-time using remotely sensed correlates of life-history parameters, a methodology already successfully demonstrated with the ticks *Rhipicephalus appendiculatus* (Randolph and Rogers, 1997) and *Boophilus decoloratus* (Estrada-Peña, 2002).

Acknowledgments

The authors wish to thank the Ministry of Health of Turkey for its collaboration and support in the elaboration of this paper. SPOT images were obtained through the courtesy of the Free Vegetation Program, copyrighted by CNES and distributed by VITO (Belgium).

References

- Bakir M, Ugurlu M, Dokuzoguz B, Bodur H, Tasyaran MA, Vahaboglu H, 2005. Crimean-Congo haemorrhagic fever outbreak in Middle Anatolia: a multicentre study of clinical features and outcome measures. *J Med Microbiol* 54, 385-389.
- Ergonul O, Celikbas A, Dokuzoguz B, Eren S, Baykam N, Esener H, 2004. The characteristics of Crimean-Congo haemorrhagic fever in a recent outbreak in Turkey and the impact of oral ribavirin therapy. *Clin Infect Dis* 39, 285-289.
- Estrada-Peña A, 2002. A simulation model of population densities, survival rates and prevalence of *Boophilus decoloratus* (Acari: Ixodidae) using remotely sensed environmental information. *Vet Parasitol* 105, 51-78.
- Estrada-Peña A, Vatanserver Z, Gargili A, Aktas M, Uzun R, Ergonul O, Jongejan F, 2007. Modelling the spatial distribution of Crimean-Congo Haemorrhagic Fever outbreaks in Turkey. *Vector Borne Zoonotic Dis* 7, in press.
- Hay SI, 1997. Remote sensing and disease control: past, present and future. *Trans R Soc Trop Med Hyg* 91, 105-106.
- Hay SI, Omumbo JA, Craig MH, Snow RW, 2000. Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. *Adv Parasitol* 47, 173-215.
- Hay SI, Snow RW, Rogers DJ, 1998. From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. *Parasitol Today* 14, 306-313.
- Hay SI, Tatem AJ, Gram AJ, Gotees SJ, Rogers DJ, 2006. Global environmental data for mapping infectious disease distribution. *Adv Parasitol* 62, 38-71.
- Hoogstraal H, 1979. The epidemiology of tick-borne Crimean-Congo haemorrhagic fever in Asia, Europe, and Africa. *J Med Entomol* 15, 307-417.
- Hugh-Jones M, 1989. Applications of remote sensing to the identification of habitats of parasites and disease vectors. *Adv Parasitol* 5, 244-251.
- Karti SS, Odabasi Z, Korten V, 2004. Crimean-Congo hemorrhagic fever in Turkey. *Emerg Infect Dis* 19, 1379-1384.
- Kristensen TK, Malone JB, McCarroll JC, 2001. Use of satellite remote sensing and geographic information systems to model the distribution and abundance of snail intermediate hosts in Africa: a preliminary model of *Biomphalaria pfeifferi* in Ethiopia. *Acta Trop* 79, 73-78.
- Malone JB, Yilma JM, McCarroll JC, Mukaratirwa S, Erko B, Zhou X, 2001. Satellite climatology and environmental risk of *Schistosoma mansoni* in Ethiopia. *Acta Trop* 79, 59-72.
- Odiit M, Bessell R, Fèvre EM, Robinson T, Kinoti J, Coleman PG, Welburn SC, McDermott J, Woolhouse MEJ, 2005. Using remote sensing and geographic information systems to identify villages at high risk for rhodesiense sleeping sickness in Uganda. *Trans R Soc Trop Med Hyg* 100, 354-362.
- Ozkurt Z, Kiki I, Erol S, 2006. Crimean-Congo hemorrhagic fever in Eastern Turkey: clinical features, risk factors and efficacy of ribavirin therapy. *J Infect* 52, 207-215.
- Perkins SE, Cattador IM, Tagliapietra V, Rizzoli AP, Hudson PJ, 2006. Localized deer absence leads to tick amplification. *Ecology* 87, 1981-1986.
- Pettorelli N, Vik JO, Mysterud A, Gaillard JM, Tucker CT, Stenseth NC, 2005. Using the satellite-derived NDVI to assess ecological response to environmental change. *Trends Ecol Evol* 20, 503-510.
- Randolph SE, Rogers DJ, 1997. A generic model for the African tick *Rhipicephalus appendiculatus*. *Parasitology*

- 115, 265-279.
- Rogers DJ, Randolph SE, 1991. Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature* 351, 739-741.
- Rogers DJ, Randolph SE, 2003. Studying the global distribution of infectious diseases using GIS and RS. *Nat Rev Microbiol* 1, 231-237.
- Rosà R, Pugliese A, Norman R, Hudson PJ, 2003. Thresholds for disease persistence in models for tick-borne infections including non-viraemic transmission, extended feeding and tick aggregation. *J Theor Biol* 224, 359-376.
- Sellers PJ, Berry JA, Collatz GJ, Field CB, Hall FG, 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sens Environ* 42, 187-216.
- Thomson MC, Connor SJ, Milligam P, Flasse SP, 1996. The ecology of malaria – as seen from Earth observation satellites. *Ann Trop Med Parasitol* 90, 243-264.
- Thomson MC, Connor SJ, Milligam P, Flasse SP, 1997. Mapping malaria risk in Africa: what can satellite data contribute? *Parasitol Today* 13, 313-318.
- Whitehouse CA, 2004. Crimean-Congo haemorrhagic fever. *Antiviral Res* 64, 145-160.