ORIGINAL PAPER

# Modelling non-equilibrium distributions of invasive species: a tale of two modelling paradigms

**R. W. Sutherst · A. S. Bourne**

**Abstract** Invasive species, biological control and climate change are driving demand for tools to estimate species' potential ranges in new environments. Flawed results from some tools are being used to inform policy and management in these fields. Independent validation of models is urgently needed so we compare the performance of the ubiquitous, logistic regression and the CLIMEX model in predicting recent range extensions of the livestock tick, *Rhipicephalus (Boophilus) microplus,* in Africa. Both models have been applied to the tick so new, independent data can be used to test their ability to model non-equilibrium distributions. Logistical regression described the spatial data well but failed to predict the range extensions. CLIMEX correctly predicted the extensions without fitting the non-equilibrium data accurately. Our results question the validity of using descriptive, statistical models to predict changes in species ranges with translocation and climate change. More test cases that include independent validation are needed.

## Introduction

The threats to biodiversity, agriculture and human health from climate change (IPCC 2007) and invasive species (Mooney and Hobbs 2000) are creating a burgeoning demand for tools to estimate changes in potential species' ranges (Holt and Keitt 2005). This demand is increased by the ongoing translocation of biological control agents (Sutherst 2003). Multivariate statistics and rule-based approaches, with environmental variables sourced from meteorological stations, ground observations or satellite imagery, have been used widely (e.g. Rogers 1995; Austin and Meyers 1996; Rogers et al. 1996; Estrada-Pena 2001; Pearson et al. 2002; Peterson 2003; Guisan and Thuiller 2007). The models are being used to make projections of the potential geographical distribution of species under current or future climates (Rogers and Randolph 1993; Beaumont and Hughes 2002; Midgley et al. 2002; Estrada-Pena et al. 2006). Such results and many others have been used to inform policy (Thomas et al. 2004), despite large differences between different models when projecting changes in ranges of the same species in response to environmental changes (Kriticos and Randall 2001;

R. W. Sutherst (✉)
Ecology Centre, School of Integrative Biology, University of Queensland, St. Lucia 4072, QLD, Australia
e-mail: r.sutherst@uq.edu.au

A. S. Bourne
Long Pocket Laboratories, CSIRO Entomology, 120 Meiers Rd, Indooroopilly 4068, QLD, Australia

Robertson et al. 2003; Thuiller 2003; Elith et al. 2006; Zeman and Lynen 2006). Rigorous validation against independent data has been lacking (Sutherst 1998) partly because opportunities for such tests with endemic species do not exist (Sutherst and Maywald 2005). On the other hand invasive species provide many such opportunities as we now demonstrate.

One invasion that allows independent validation is provided by the pan-tropical bovine tick, *Rhipicephalus (Boophilus) microplus* (Canestrini) (Acari:Ixididae), which has been expanding its range in Africa over a period of decades. We compare *a priori* estimates of the areas at risk of further invasion in Africa based on a multivariate logistic regression model (Cumming 1999a, b; Cumming 2002) and the inference-based, climatic response model, CLIMEX, (Sutherst and Maywald 1985; Sutherst et al. 1995; Sutherst 2003).

### Historical geographical distribution of *R. B. microplus*

*R. B. microplus* is a bovine tick of Asian origin that is now widespread in tropical and sub-tropical Central and South America, Australia and the east coast of Africa (Wharton 1974; Cumming 1999a, b). It has been in south-east Africa for over a century (Theiler 1949), during which time its distribution has waxed and waned in southern Africa prior to 1999 as cited by Lynen et al. (2008). Theiler (1943) foresaw future extensions of the localised foci in southern Africa. The prior occupation of much of the continent by the very similar endemic species, *B. decoloratus*, which mates with *R. B. microplus* to produce sterile hybrid zones (Sutherst 1987a) has been a barrier to the latter's spread. Modelling suggested that it was possible for *R. B. microplus* to colonise the occupied territory in much of sub-Saharan Africa under certain conditions of climate, host suitability and propagule pressure. As such this exotic tick, which transmits *Babesia bovis*, a more pathogenic bovine parasite than the endemic *B. bigemina*, is a potential threat to bovines in other parts of Africa.

Recently *R. B. microplus* expanded its range slightly into marginal country at the northernmost tip of South Africa during a period of above average rainfall (Tønnesen et al. 2004) and northern Tanzania (Lynen et al. 2008). Most recently the tick has been recorded for the first time from a 'dense humid forest' habitat 50 km north of Abidjan in Ivory Coast, West Africa (Madder et al. 2007), as shown with stars in Figs. 1b and 2. These invasions provide an excellent opportunity to compare model predictions using truly independent observations.

### Model comparison

A regression model with up to 49 predictor variables (mean monthly rainfall, maximum and minimum temperatures, normalized difference vegetation index (NDVI) and elevation) was built for 73 species of ticks to describe Pan-African distributions of species richness (Cumming 1999a, b; Cumming 2002). The climatic data used were restricted to the southern hemisphere. A number of studies were cited (Cumming 1999a, b) that used multivariate analyses to estimate the likely range of particular species, and to infer the role of different environmental variables in determining species' distributions. Some of the limitations of descriptive methods were pointed out but these caveats have not been taken into account by later users of regression and other statistical methods.

The observed and estimated distributions of *R. B. microplus* in Africa (Cumming 1999a, b) are shown in Fig. 1a, b. The model provides a very credible description of the data. It also shows small areas at risk in Namibia and Angola. Zero or minimal risk is attributed to most of East and Central Africa. The results were reproduced in a series of high profile maps by a consortium supported by the UK Department for International Development's (DFID) Animal Health Programme, Food and Agriculture Organization (FAO), International Livestock Research Institute (ILRI) and the Integrated Consortium on Ticks and Tick-borne Diseases (ICTTD), with tick and tick-borne disease interests in Africa (Minjauw et al. 2001).

We repeated Cumming's analysis, using only climatic variables because Cumming stated that climate accounted for nearly all of the observed variation, due to the high correlation between the predictive variables. The presence of hosts and vegetation types had also been eliminated as limiting factors (Cumming 1999b, 2002). We ran the model using an interpolated climate surface (monthly maximum and minimum temperature, rainfall and relative humidity) with a resolution of 10′ (New et al. 2002). Our results were sufficiently similar to those of
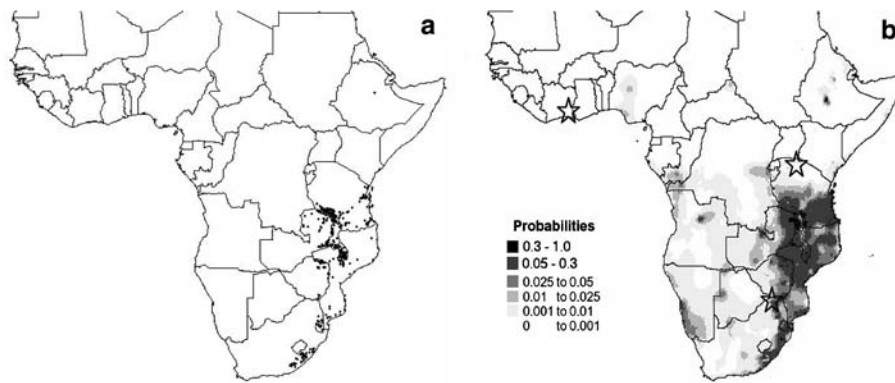
Fig. 1 (a) Observed and (b) estimated potential geographical distribution of the bovine tick, *R. B. microplus* in Africa, using logistic regression fitted to the observed African distribution.

The stars show the areas in Venda, north-western Tanzania and Ivory Coast in West Africa from where the tick has recently been recorded

(Cumming 1999a, b) to ensure that a climate-based model comparison was valid for further comparisons. A question arose as to what weighting effect the different density of observational records had on the estimation of the model parameter values. The data are heavily biased towards northern Malawi and south-western Tanzania. This coincides with the area with greatest predicted habitat suitability. We tested the effect of the sampling bias, by aggregating clusters of data points, and found that it had little discernable effect on the predicted areas. A major effect on the range limits arose from the weighting given to absence records.

The potential geographical distribution of *R. B. microplus* in relation to climate had also been estimated for Africa (Sutherst and Maywald 1985; Sutherst 1987b), using the CLIMEX model. CLI-MEX is a simplified simulation model that infers the responses of species to average maximum and minimum temperatures, rainfall and relative humidity from its geographical distribution and seasonal phenology. It incorporates a hydrological model and runs on a weekly time step. The model combines growth and stress parameters to produce a measure—the Ecoclimatic Index (EI)—of the overall suitability of the climate for the target species. The original CLIMEX model of *R. B. microplus* was parameterised using the distributions in Australia (Sutherst and Maywald 1985) and the south coast of South Africa because a quarantine boundary prevented southward spread beyond the Queensland, New South Wales

border, so the cold-limited boundary could not be estimated precisely in Australia (Sutherst 1987a, b).

In order to make the model comparison fair and equal, we re-fitted the CLIMEX parameters using only the distribution records of Cumming (1999a, b) (Fig. 1) and the results are shown in Fig. 2. There was minimal difference between these results and those of Sutherst and Maywald (1985) and those differences related to the cold limits only which do not affect the current argument. CLIMEX was unable to include the observed records of occurrence without also including a large additional area beyond the northern limits of the data. The map depicts the suitability of the climate for the tick in the whole of Africa and hence the area predicted to be at risk of invasion in the absence of other constraints. It indicates that the highest risk areas in Africa have not yet been invaded, i.e. in East, Central Africa including most of Angola and West Africa.

Both models identified Venda as being moderately suitable. CLIMEX also identified north-western Tanzania and Ivory Coast in West Africa as being highly favourable for *R. B. microplus*. In addition, it flagged the high risk to much of the highlands of East and Central Africa, and the humid tropics of West Africa. Further, it revealed a western corridor of suitable habitat connecting southern and north-western Tanzania. The regression model classified northern Tanzania as unsuitable and it could not be applied to West Africa because it was parameterised with climatic data from the southern hemisphere.

**Fig. 2** The potential
geographical distribution
of the bovine tick,
*R. B. microplus* in Africa,
estimated using the
CLIMEX model and the
Australian geographical
distribution. The stars show
the areas in Venda, north-
western Tanzania and Ivory
Coast in West Africa from
where the tick has recently
been recorded
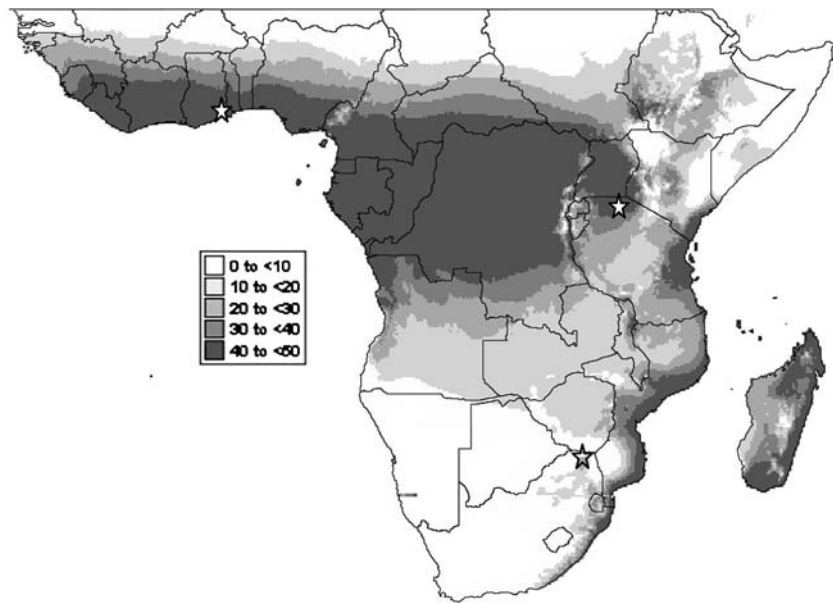


## Discussion

Comparison of the results from the two models
(Figs. 1b, 2) could not provide a more stark contrast
in prediction of invasive risk. The regression
described the observed distribution records of
*R. B. microplus* in southern Africa accurately but
excluded the main range extensions. On the other
hand the CLIMEX results included the observed
distribution but also included large areas outside the
range of the observations, which included the
observed range extensions. The reasons for these
differences go to the heart of modelling non-equilib-
rium populations.

Firstly, the regression algorithm and parameter
fitting process treat each data point as a presence or
absence record. This has the unavoidable effect of
categorising those areas that the species has not yet
colonized as unsuitable. Such a model is therefore
inherently unsuitable for anticipating range exten-
sions of invasive species that have not yet reached
equilibrium in their new environment. With CLI-
MEX, absence data—in areas that are estimated to be
suitable from the climatic conditions associated with
presence data—prevent the model from restricting its
parameter values to simulating only the presence
records. Such absence records are attributed
'unknown' status in such conditions. They stimulate
a search for other explanations.

Secondly, the regression fitting process is limited
to using truncated or selective independent variables.
In the current application, despite the use of numer-
ous climatic variables, the model is unable to take
account of the different seasonal patterns across
latitudes and longitudes. While southern Africa has
mostly a summer rainfall pattern, further north the
rainfall seasonality is quite different. The pattern-
matching algorithm categorises different combina-
tions of climatic variables as having low suitability
because they do not correspond well with the patterns
associated with the observations.

The inability of statistical models to accommodate
different patterns of environmental variables has been
noted previously (Kriticos and Randall 2001). This
failure can be partly overcome by choosing variables
that are more independent of the structure of seasonal
cycles of temperature and rainfall. Annual average
temperatures and rainfall (Rogers and Randolph
2000) or derived variables like annual degree-days
or moisture indices based on ratios of rainfall to
evaporation (Austin and Meyers 1996; Lindsay et al.
1998) have been used. However they do not take
account of the needs of species to have concurrent
combinations of suitable temperature and moisture
(Sutherst and Maywald 2005). They also only
estimate conditions suitable for growth and do not
take account of the limiting effects of extreme
conditions of different durations. The result is risk

assessments with vast scope for bias and error from ignoring both seasonality and interactions between temperature and moisture variables.

In discussing the major differences between the original CLIMEX model predictions of the area at risk in Africa and the historical observed distribution, the authors stated that 'Clearly areas such as western Kenya have a high risk from *R. B. microplus* if that species should be introduced there'. They also stated that 'discrepancies between observed and predicted distributions can be just as useful as agreements in identifying limiting factors' (Sutherst and Maywald 1985). In the case of *R. B. microplus* this discrepancy contributed to identifying the hybrid zone phenomenon involving this tick and *B. decoloratus* in Africa (Sutherst 1987a, b).

In fact, the reported distribution of *R. B. microplus* in southern Africa includes almost all of the range of temperature and moisture values that define the tick's climatic domain. The tick has a preference for a warm, humid climate as found in tropical Asia and also in East, Central and West Africa. The unsuitable rating from the regression model was produced even though the most basic consideration of the temperature and moisture requirements of the tick indicates that the climates are highly favourable. Indeed other similar *Boophilus* species occupy large parts of sub-Saharan Africa (Hoogstraal 1956; Cumming 1999a, b).

The major differences between the results from using regression compared with CLIMEX arise in part from the different objectives of the different methods. Firstly, statistical models often used in conjunction with geographical information systems (GIS) are powerful tools for interpolating between sample data to fill in missing values. They are not appropriate for extrapolating beyond the data sets as is necessary with species invasions or climate change scenarios. They can describe geographical distributions precisely (Rogers and Randolph 1993) but the dangerous assumption is often made that the models therefore enable valid extrapolation to other regions or global change scenarios. The better the fit to the data the more successful the model is considered to be. The inappropriateness of using such models to describe species distributions of tsetse fly (*Glossina* spp.) (Rogers and Packer 1993; Rogers 1995; Robinson et al. 1997), ticks (Rogers and Randolph 1993) and malaria (Rogers and Randolph 2000) has been alluded to previously (Sutherst 1998, 2004).

Many claims of high accuracy of statistical and simulation models in describing species' potential distributions arise from comparison of highly correlated subsets of data obtained by splitting datasets for training and validation exercises. This inappropriate practice is widespread and may have arisen from a misunderstanding of the recommendation to split data sets in order to obtain unbiased estimates of parameter values after selecting models using the original subset (Miller 2002).

These findings show that using regression, and by implication any descriptive pattern-matching approach, is not appropriate when predicting potential range changes under global change or invasion scenarios. The failure of such models to make provision for non-equilibrium ranges, as shown with this example of *R. B. microplus* in Africa demonstrates their weaknesses wherever extrapolations are required. This has led to the misguided practice of using ensembles of statistical models (Thuiller 2003) in an attempt to reduce the errors. That does not address the underlying problem of inappropriate model structures and optimisation methods. CLIMEX on the other hand aims to infer the species climatic response relationships from the known geographical distribution and then to project likely responses to climates in different places and climate change scenarios. Rather than trying to achieve a precise description of the distribution, the emphasis is on interrogating the data to shed light on the climatic conditions that support growth or limit the survival of the species. The need to ensure internal consistency in the results is paramount and the aim is to include all known positive locality records. Inconsistencies, like those above, warn that climate alone is not responsible for limiting the geographical distribution, or that the species has not reached its equilibrium distribution, rather than indicating model failure.

Most native species have spread to their maximum potential ranges and so do not provide the opportunity for an a priori assessment of areas suitable for further spread. An exception was the invasion of the highlands of Zimbabwe by two African species of ticks, *Amblyomma variegatum* and *Amblyomma hebraeum* (Bruce and Wilson 1998). This occurred after a warning of such a risk was given a decade earlier based on a CLIMEX analysis (Sutherst and Maywald 1985). Such apparent discontinuities or inconsistencies in geographical distributions may also arise from

incomplete distribution records and CLIMEX can help to identify such gaps (Hall and Wall 1995; Sutherst 2003; Sutherst and Maywald 2005). *R. B. microplus* can be expected to continue its spread through the continent and occupy those habitats to which it is better adapted than the resident *Boophilus* species. The continuing spread will provide further opportunities for model validation.

We repeat the appeal (Sutherst and Maywald 2005) for less concern with precision of model fits to species geographical distributions (Cumming 2000a, b; Anderson and Lew 2003; Elith et al. 2006), which are usually of poor quality and more emphasis on avoiding the systemic errors that arise from faulty model structures. Such errors are not revealed by statistical measures of goodness of fit. We now add the issue of inappropriate selection of predictive environmental variables and weighting given to absence records. Derived measures of climatic suitability may improve the performance of statistical models but will not enable them to match the performance of dynamic simulation models. The mechanistic CLIMEX simulation model performed reliably using such derived variables and provides more insight into the role of specific growth and limiting conditions across all global climates. As such, it is a useful first step in an ecological study of any species (Sutherst et al. 2007). Being mechanistic, each parameterised function constitutes a testable hypothesis on the role that the climatic factor has in supporting population growth or limiting a species range (Sutherst et al. 1995). The relationships can inform policy and guide further process-based studies.

As ecologists we need to go back to the basics of our discipline and consider the needs of species from first principles ('to look at the world through the eyes of the species we are studying') instead of applying descriptive models to multiple species using a black-box 'plug-and-play' approach as done for many studies such as those cited by (Thomas et al. 2004). Risk assessments for biodiversity under global change will be better served by selecting species with ranges that are representative of their compatriots and understanding them thoroughly as a basis for extrapolation to the other species with similar ranges. Progress in understanding the factors limiting species' geographical distributions will only come from applying mechanistic, hypothesis-based approaches to the analyses. This, after all, is the fundamental

distinguishing characteristic of science and it leaves little room for descriptive methods.

## References

Anderson RP, Lew D (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Modell 162:211–232

Austin MP, Meyers JA (1996) Current approaches to modelling the environment niche of eucalypts: implication for management of forest biodiversity. For Ecol Manag 85:95–106

Beaumont LJ, Hughes L (2002) Potential changes in the distribution of latitudenally restricted Australian butterfly species in response to climate change. Glob Chang Biol 8:954–971

Bruce D, Wilson A (1998) The extensive spread of heartwater in Zimbabwe since 1975. In: Coons L, Rothschild M (eds) The second international conference on tick-borne pathogens at the host-vector interface: a global perspective, August 28–September 1 1995, Kruger National Park, South Africa, S.I. The Conference, pp 105–112

Cumming GS (1999a) The evolutionary ecology of African ticks. Oxford University, Oxford

Cumming GS (1999b) Host distributions do not limit the species ranges of most African ticks (Acari: Ixodida [Metastigmata]). Bull Entomol Res 89:303–327

Cumming GS (2000a) Using between-model comparisons to fine-tune linear models of species ranges. J Biogeogr 27:441–455

Cumming GS (2000b) Using habitat models to map diversity: Pan-African species richness of ticks (Acari: Ixodida). J Biogeogr 27:425–440

Cumming GS (2002) Comparing climate and vegetation as limiting factors for species ranges of African ticks. Ecology 83:255–268

Elith J, Graham CH et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29(2):129–151

Estrada-Pena A (2001) Climate warming and changes in habitat suitability for *Boophilus microplus* (Acari: Ixodidae) in Central America. J Parasitol 87(5):978–987

Estrada-Pena A, Corson M et al (2006) Changes in climate and habitat suitability for the cattle tick *Boophilus microplus* in its southern Neotropical distribution range. J Vector Ecol 31(1):158–167

Guisan A, Thuiller W (2007) Predicting species distribution: offering more than simple habitat models (vol 8, pg 993, 2005). Ecol Lett 10(5):435

Hall M, Wall R (1995) Myiasis of humans and domestic animals. Adv Parasitol 35:257–334

Holt RD, Keitt TH (2005) Species' borders: a unifying theme in ecology. Oikos 108(1):3–6

Hoogstraal H (1956) Ticks of the Sudan. US NAMRU No 3, Cairo

IPCC (2007) Climate change 2007: synthesis report. Cambridge University Press

Kriticos DJ, Randall PR (2001) A comparison of systems to analyse potential weed distributions. In: Groves RH, Panetta FD, Virtue JG (eds) Weed risk assessment. CSIRO, Melbourne, pp 61–79

Lindsay SW, Parson L et al (1998) Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data. Proc R Soc Lond B 265: 847–853

Lynen G, Zeman P et al (2008) Shifts in the distributional ranges of *Boophilus* ticks in Tanzania: evidence that a parapatric boundary between *Boophilus microplus* and *B. decoloratus* follows climate gradients. Exp Appl Acarol 44:147–164

Madder M, Thys E et al (2007) *Boophilus microplus* ticks found in West Africa. Exp Appl Acarol 43:233–234

Midgley GF, Hannah L et al (2002) Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. Glob Ecol Biogeogr 11:445–451

Miller A (2002) Subset selection in regression. CRC Press (Chapman & Hall), London

Minjauw B, Robinson T et al (2001) Map of *Anaplasmosis* and *Babesiosis* in sub-Saharan Africa. ILRI, Nairobi

Mooney HA, Hobbs RJ (2000) Invasive species in a changing world. Island Press, Corelo

New M, Lister D et al (2002) A high-resolution data set of surface climate over global land areas. Clim Res 21:1–25

Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. Q Rev Biol 78:419–433

Pearson RG, Dawson TP et al (2002) SPECIES: a spatial evaluation of climate impact on the envelope of species. Ecol Modell 154:289–300

Robertson MP, Peter CI et al (2003) Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modelling techniques. Ecol Modell 16:153–167

Robinson TP, Rogers DJ et al (1997) Mapping tsetse habitat suitability in the common fly belt of southern Africa using multivariate analysis of climate and remotely sensed vegetation data. Med Vet Entomol 11:235–245

Rogers DJ (1995) Remote sensing and the changing distribution of tsetse flies in Africa. In: Harrington R, Stork NE (eds) Insects in a changing environment. Academic Press, London, pp 177–193

Rogers DJ, Packer MJ (1993) Vector-borne diseases, models, and global change. Lancet 342:1282–1284

Rogers KJ, Randolph SE (1993) Distribution of tsetse and ticks in Africa: past, present and future. Parasitol Today 9(7):266–271

Rogers DJ, Randolph SE (2000) The global spread of malaria in a future, warmer world. Science 289:1763–1766

Rogers DJ, Hay SI et al (1996) Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. Ann Trop Med Parasitol 90(3):225–241

Sutherst RW (1987a) The dynamics of hybrid zones between tick (Acari) species. Int J Parasitol 17:921–926

Sutherst RW (1987b) The role of models in tick control. In: Hughes KL (ed) Proceedings of the international conference on veterinary preventive medicine and animal production. Australian Veterinary Association, Melbourne, pp 32–37

Sutherst RW (1998) Implications of global change and climate variability for vector-borne diseases: generic approaches to impact assessments. Int J Parasitol 28:935–945

Sutherst RW (2003) Prediction of species geographical ranges. Guest editorial. J Biogeogr 30:1–12

Sutherst RW (2004) Global change and human vulnerability to vector-borne diseases. Clin Microbiol Rev 17(1):136–173

Sutherst RW, Maywald GF (1985) A computerised system for matching climates in ecology. Agric Ecosyst Environ 13(3–4):281–300

Sutherst RW, Maywald G (2005) A climate-model of the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae): implications for invasion of new regions, particularly Oceania. Environ Entomol 34:317–335

Sutherst RW, Maywald GF et al (1995) Predicting insect distributions in a changing climate. In: Harrington R, Stork NE (eds) Insects in a changing environment. Academic Press, London, pp 59–91

Sutherst RW, Maywald GF et al (2007) Including species-interactions in risk assessments for global change. Glob Chang Biol 13(9):1843–1859

Theiler G (1943) "Notes on the ticks off domestic stock from Portuguese East Africa." Estação anti-malárica de Lourenço Marques: 1–55

Theiler G (1949) Zoological survey of the union of South Africa. Tick survey: part II. Distribution of *Boophilus (Palpoboophilus) decoloratus*, the blue tick. Onderstepoort J Vet Sci Anim Ind 22:255–268

Thomas D, Cameron A et al (2004) Extinction risk from climate change. Nature 427:145–148

Thuiller W (2003) Biomod-optimizing predictions of species distributions and projecting potential future shifts under global change. Glob Chang Biol 9(10):1353–1362

Tønnesen MH, Penzhorn BL et al (2004) Displacement of *Boophilus decoloratus* by *Boophilus microplus* in the Soutpansberg Region, Limpopo Province, South Africa. Exp Appl Acarol 32:199–208

Wharton RH (1974) The current status and prospects for the control of ixodid ticks with special emphasis on *Boophilus microplus*. Bulletin de l'Office International des Epizooties 81:65–85

Zeman P, Lynen G (2006) Evaluation of four modelling techniques to predict the potential distribution of ticks using indigenous cattle infestations as calibration data. Exp Appl Acarol 39:163–176