Predicting the potential environmental suitability for *Theileria orientalis* transmission in New Zealand cattle using maximum entropy niche modelling

K.E. Lawrence [a],[*], S.R. Summers [a], A.C.G. Heath [b], A.M.J. McFadden [c], D.J. Pulford [c], W.E. Pomroy [a]

[a] Massey University, Palmerston North, New Zealand
[b] Agresearch Ltd., c/o Hopkirk Research Institute, Private Bag 11008, Palmerston North 4442, New Zealand
[c] Ministry for Primary Industries, PO Box 40742, Upper Hutt 5018, New Zealand

ARTICLE INFO

ABSTRACT

The tick-borne haemoparasite *Theileria orientalis* is the most important infectious cause of anaemia in New Zealand cattle. Since 2012 a previously unrecorded type, *T. orientalis* type 2 (Ikeda), has been associated with disease outbreaks of anaemia, lethargy, jaundice and deaths on over 1000 New Zealand cattle farms, with most of the affected farms found in the upper North Island. The aim of this study was to model the relative environmental suitability for *T. orientalis* transmission throughout New Zealand, to predict the proportion of cattle farms potentially suitable for active *T. orientalis* infection by region, island and the whole of New Zealand and to estimate the average relative environmental suitability per farm by region, island and the whole of New Zealand.

The relative environmental suitability for *T. orientalis* transmission was estimated using the Maxent (maximum entropy) modelling program. The Maxent model predicted that 99% of North Island cattle farms (n = 36,257), 64% South Island cattle farms (n = 15,542) and 89% of New Zealand cattle farms overall (n = 51,799) could potentially be suitable for *T. orientalis* transmission. The average relative environmental suitability of *T. orientalis* transmission at the farm level was 0.34 in the North Island, 0.02 in the South Island and 0.24 overall.

The study showed that the potential spatial distribution of *T. orientalis* environmental suitability was much greater than presumed in the early part of the *Theileria* associated bovine anaemia (TABA) epidemic. Maximum entropy offers a computer efficient method of modelling the probability of habitat suitability for an arthropod vectored disease. This model could help estimate the boundaries of the endemically stable and endemically unstable areas for *T. orientalis* transmission within New Zealand and be of considerable value in informing practitioner and farmer biosecurity decisions in these respective areas.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Theileria orientalis is a tick-borne obligate intracellular protozoan parasite of cattle and buffaloes and belongs to the phylum Apicomplexa (Watts et al., 2016). This phylum includes some of the most pathogenic parasitic protozoa of humans and domestic animals (Morrissette and Sibley, 2002). Eleven different *T. orientalis* genotypes have been described including Type 1 (Chitose), Type 2 (Ikeda) and Type 3 (Buffeli). Types 4–8 and Types N1-N3 which remain unnamed (Khukhuu et al., 2011; Kamau et al., 2011; Cufos et al., 2012; Watts et al., 2016; Hammer et al., 2015). Infected cattle sometimes develop *Theileria*-associated bovine anaemia (TABA) which can be a severe and life threatening (McFadden et al., 2011). Some genotypes are more pathogenic than others.

There have been two recorded epidemics of TABA in New Zealand, the first starting in 1982, peaking at 60 cases in 1985 before almost completely disappearing after 1986 (James et al., 1984; Thompson, 1991). The second outbreak started in August 2012 and is still ongoing (McFadden et al., 2013; Lawrence et al.,

* Corresponding author.
E-mail addresses: K.Lawrence@massey.ac.nz (K.E. Lawrence),
S.R.Summers@massey.ac.nz (S.R. Summers), acgandmm.heath@actrix.co.nz
(A.C.G. Heath), Andrew.McFadden@mpi.govt.nz (A.M.J. McFadden),
David.Pulford@mpi.govt.nz (D.J. Pulford), W.Pomroy@massey.ac.nz (W.E. Pomroy).
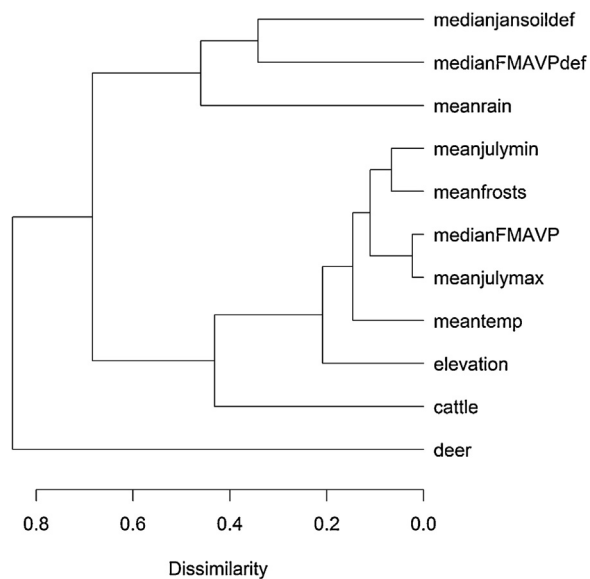
Fig. 1. Dendrogram showing clustering of environmental variables, the horizontal axis shows the dissimilarity distance between the variables.

2016), with over 1000 cattle farms since August 2012, mostly in the upper North Island, having at least one recorded clinical case of TABA. Molecular investigation of sporadic outbreaks of TABA since 1982 but prior to the current 2012 outbreak showed that the first TABA outbreak in New Zealand was probably associated with the *T. orientalis* types, Type 1 (Chitose) and Type 3 (Buffeli) (McFadden et al., 2011). The present epidemic is strongly associated with Type 2 (Ikeda) which had been previously unrecorded in New Zealand (Gias et al., 2016).

The Ikeda type is much more pathogenic than either the Chitose or Buffeli types (Hammer et al., 2015). Even so, for most New Zealand cattle farms Ikeda infection is not associated with high incidence of clinical disease as the median mortality and morbidity rates in newly infected cattle herds are reported as 0.23% and 0.97% respectively (Vink et al., 2016). However, there have also been a small number of herds where the mortality rate has exceeded 5%, often without any clear indication why this has occurred (McFadden et al., 2013). Many infected animals are sub-clinical (Fairley, 1992) but if clinical disease does develop the following clinical signs have been observed: anaemia, abortion, lethargy, production drop, anorexia, ill-thrift, constipation, jaundice, collapse, pale udders, haemoglobinuria (less commonly) and inappetance (McFadden et al., 2013). On dairy farms, mature cows are most often affected around calving or in early lactation, i.e. at peak stress periods, whilst on beef herds calves between the age of 2 and 4 months are more usually affected (Lawrence et al., 2016).

The *T. orientalis* group is widely disseminated around the world due to the favourable distribution of suitable tick vectors in the genus *Haemaphysalis*. In New Zealand *H. longicornis* is the only known competent vector for *T. orientalis*. This tick species is widespread in the North Island, occurring as far south as Waikanae in the west and the Wairarapa in the east, it is also established in the top of the South Island (Heath, 2013, 2016). In New Zealand the *H. longicornis* life cycle is usually completed within 12 months, with over-wintering nymphs mainly engorging from July to September, adults from November to December and larvae from February to April (Fig. 2, Heath, 2016). Once infected, cattle probably remain infected and infectious for life (Onuma et al., 1998) and although biting flies, sucking lice, mosquitoes and iatrogenic transmission have been suggested as additional mechanical vectors of *T. orien-*

talis, there is little field data to support these claims (Fujusaki et al., 1993; Hadi and Al-Amery, 2012; Hammer et al., 2015).

Environmental niche modelling, also known as species distribution modelling, is defined as the quantitative modelling of species-environment relationships in order to interpolate species habitat suitability over much larger spatial extents (Franklin, 2010). The classic approach to niche modelling condenses the biotic (B), the abiotic (A) and the mobility (M) requirements for a species into a heuristic framework known as a Biotic Abiotic Mobility (BAM) diagram (Peterson, 2014). This is usually portrayed as a 3 circle Venn diagram, where each circle is either the B, A or M requirement and the realised niche for a species is described by the intersection B ∩ A ∩ M.

Since *T. orientalis* is an obligate intracellular parasite its geographic distribution is potentially equivalent to that of its reservoir host, cattle. However active transmission of *T. orientalis* infection between infected and naïve cattle will only occur where the distribution of cattle intersects with the realised niche for *H. longicornis* ticks. Incorporating these assumptions into a single heuristic framework for modelling purposes means that the intersection $B_{tick} \cap A_{tick} \cap M_{cattle}$ should describe the realised niche for *T. orientalis* transmission in New Zealand. It is important to acknowledge that this framework does make the assumption that the environmental requirements for the development of *T. orientalis* within the tick are similar to those required by the tick for survival off host.

Entropy means uncertainty and a probability distribution is said to have maximum entropy when it is closest to uniform (most spread out) but still explains any constraints in the data (Phillips et al., 2006). The three aims of this paper are to use maximum entropy (Maxent) to predict the relative environmental suitability for *T. orientalis* transmission across New Zealand, to estimate the proportion of cattle farms with potentially suitable environments for active *T. orientalis* transmission and to estimate the average farm relative environmental suitability by region, island and the whole of New Zealand.

2. Materials and methods

2.1. Maxent program

In this paper a machine-learning approach was used to map the probability of environmental suitability for *T. orientalis* transmission using the Maxent modelling program (Phillips et al., 2006; Phillips and Dudík, 2008), adapting the methodology described by Slater and Michael (2012). Version 3.3.3k of the Maxent maximum entropy species distribution model was used in this analysis.

Since the true prevalence of TABA disease is unknown and the case data were collected in a non-random fashion, a Bayesian solution, although attractive from a modelling perspective, is not definable (Elith et al., 2011). Maxent provides an elegant solution to this impasse by fitting a probability distribution to the data using the principle of maximum entropy. This involves fitting a probability distribution to the multivariate environmental data at the case farms, with the only constraint being that it should have the same mean as the case data, and then contrasting it with a uniform probability distribution fitted to the background sampling points. Maxent then selects the probability distribution for the case environmental data that minimises the entropy between these two distributions, which has the effect of maximising the entropy of the final probability distribution (Elith et al., 2011). This probability distribution is then used to predict the relative environmental suitability for every 500 m × 500 m pixel across a raster map of New Zealand. A raster map is an efficient method of storing spatial data as an array of equal sized cells (pixels) where each pixel contains a value. Using the default logistic output from Maxent a raster map of the relative