



Modeling the spatial distribution of *Theileria parva* (Theiler 1904), causative agent of East Coast Fever disease in cattle

Elia Vajana¹, Licia Colli¹, Marco Milanesi¹, Lorenzo Bomba¹, Riccardo Negrini^{1, 2}, Stefano Capomaccio¹, Elisa Eufemi¹, Raffaele Mazza², Alessandra Stella³, Stéphane Joost⁴, Sylvie Stucki⁴, Pierre Taberlet⁵, François Pompanon⁵, Fred Kabi⁶, Vincent Muwanika⁶, Charles Masembe⁶, Paolo Ajmone-Marsan¹, The NEXTGEN Consortium⁷

(1) Istituto di Zootecnica e Centro di Ricerca BioDNA, Facoltà di Scienze Agrarie, Alimentari ed Ambientali, Università Cattolica del S. Cuore, Piacenza, Italy, (2) AIA - Associazione Italiana Allevatori, Rome, Italy, (3) IBBA-CNR and FPTP - Fondazione Parco Tecnologico Padano, Lodi, Italy, (4) Geographic Information Systems Laboratory, EPFL - LASIG, Station 18, CH-1015 Lausanne, Switzerland, (5) LECA - Lab. d'Ecologie Alpine, UJF-CNRS, Grenoble, France, (6) Institute of Environment & Natural Resources, Makerere University, Kampala, Uganda, (7) EU funded project, <http://nextgen.epfl.ch>

Introduction

Theileria parva is a protozoan emu-parasite affecting sub-Saharan *Bos taurus* and *Bos indicus* populations. It is the causative agent of East Coast Fever, a major cattle disease causing the death of $\sim 1.1 \cdot 10^6$ animals per year and an annual loss of $\sim 168 \cdot 10^6$ USD (Norval *et al.*, 1992). *T. parva* geographical occurrence is bound to the presence of susceptible bovine host populations, the main tick vector *Rhipicephalus appendiculatus* (Neumann 1901), as well as suitable ecological conditions for the survival of both the vector and the parasite. While tick habitat requirements have been extensively investigated (see e.g. Cumming, 2002), studies focusing solely on the environmental conditions determining the parasite occurrence are still lacking.

Aim of the research: To define *T. parva* ecological fundamental niche (*sensu* Peterson & Soberón, 2012), thus fostering our understanding of the environmental requirements needed to maintain the parasite-vector-host biological system.

Materials & Methods

In the course of the EU-funded project Nextgen, 590 cattle blood samples from 204 locations covering the whole Ugandan country (Fig.1A) have been tested for the presence/absence of *T. parva* DNA (Fig 1B). For each sampling site, the values of 19 bioclimatic variables were derived from the WorldClim database (Version 1.4, release 3, <http://www.worldclim.org/>). Bioclimatic predictors and samples coordinates have been tested using *i*) a classification tree model approach and *ii*) univariate and bivariate logistic regression mixed-effects models (Eqn. 1). Bivariate models were run on significant variables from the univariate analysis and for only those pairs whose correlation didn't exceed the threshold of 0.7 (Dorman *et al.*, 2013). Statistical analyses were performed in R (Bates *et al.*, 2014; Ripley, 2014; Therneau *et al.*, 2014).

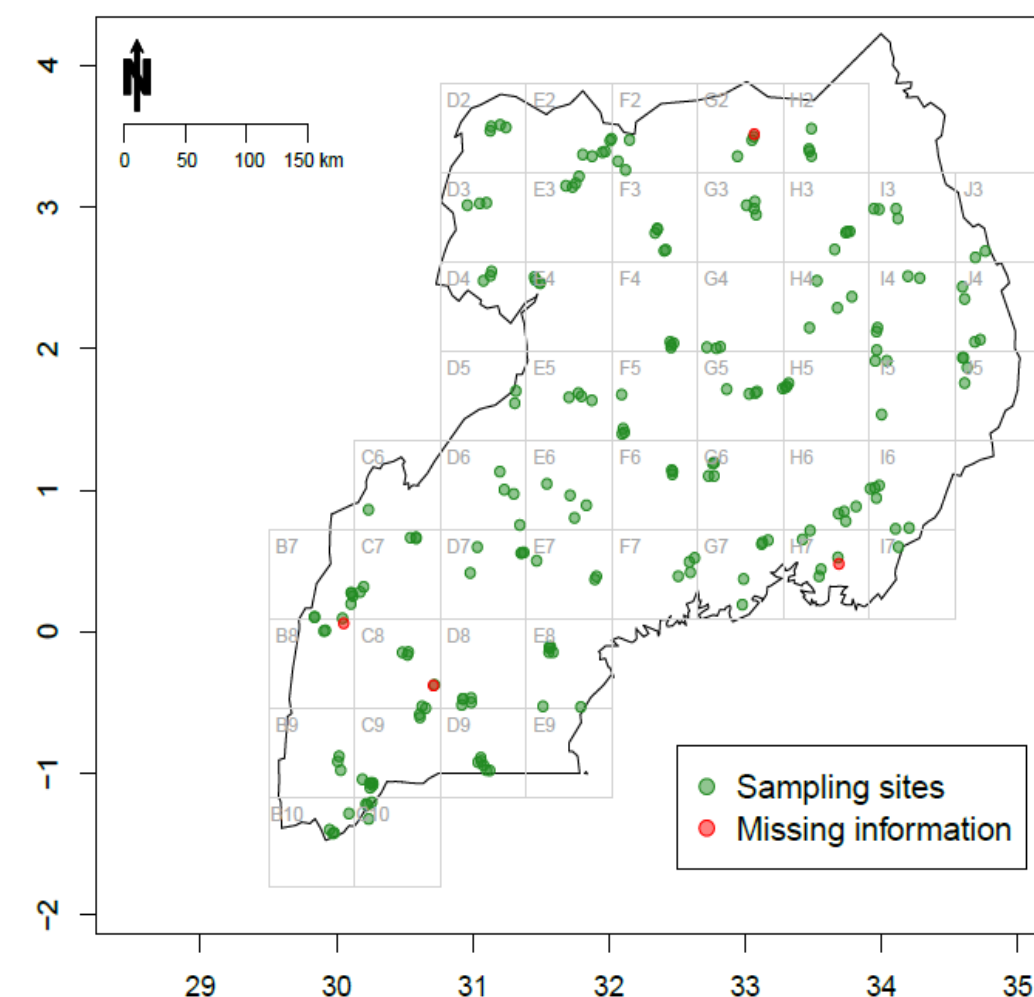


Fig. 1A – Sampling sites in Uganda.

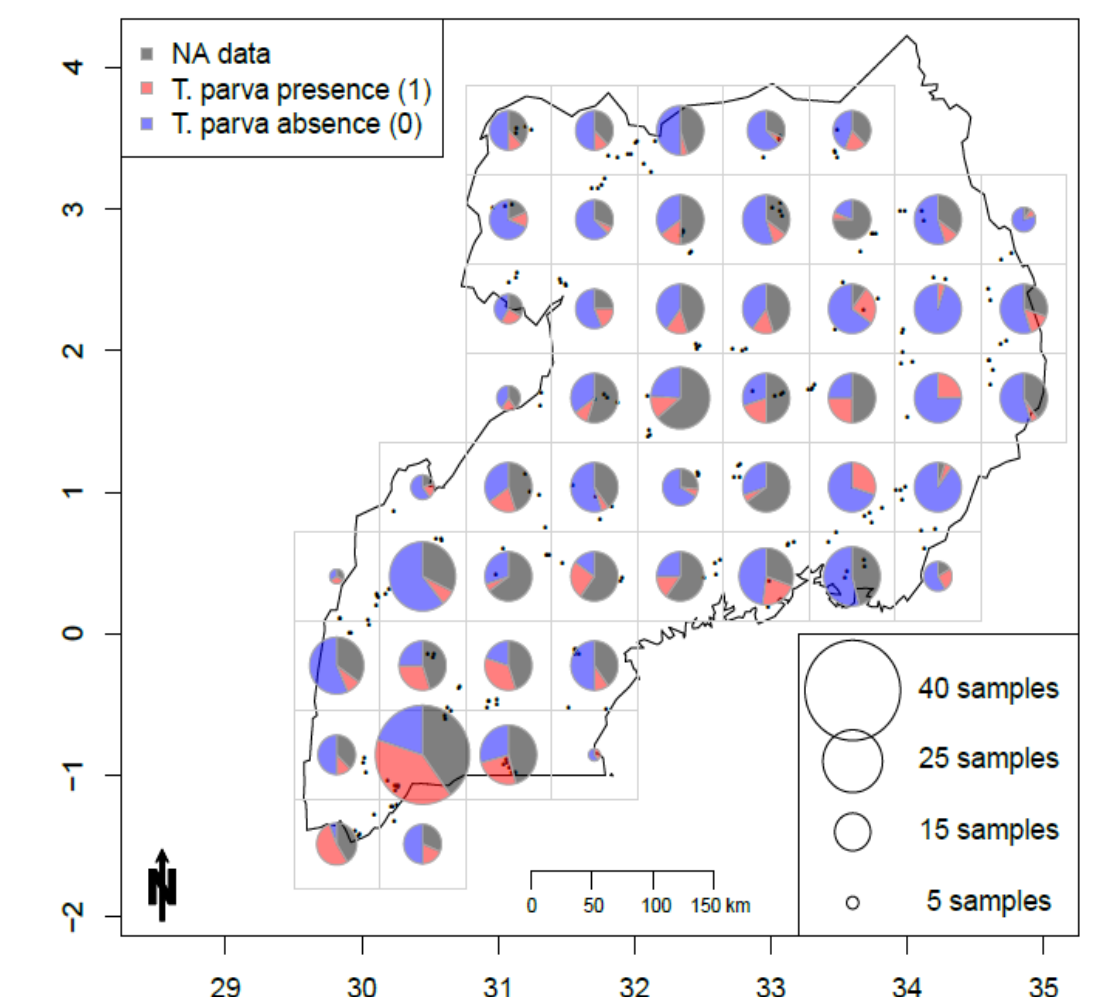


Fig. 1B – *Theileria parva* occurrence (red slice of the pie) and incidence of missing data (gray slice) per grid cell.

Results

- Classification tree and regression models highlight an overall role of latitude and return common variables related to temperature variability (Temperature seasonality, **bio4**, and annual range, **bio7**) (Fig. 3 and Tab. 1). More stable areas are associated with higher probabilities of *T. parva* presence (bio4 and bio7 models Fig. 4).
- While some associations likely reflect the vector environmental requirements (see e.g. negative relationship with **bio9**, Mean Temperature of Driest Quarter, and positive with **bio18**, Precipitation of Warmest Quarter – not shown here –), some others may represent signals of *T. parva* specific environmental needs (see negative association with **bio8**, Mean Temperature of Wettest Quarter).
- Regression models either show *i*) a probability increase following a north-west / south-east direction, or *ii*) point out mountain regions of the country, thus highlighting the existence of a not yet recognized effect (Fig. 4).

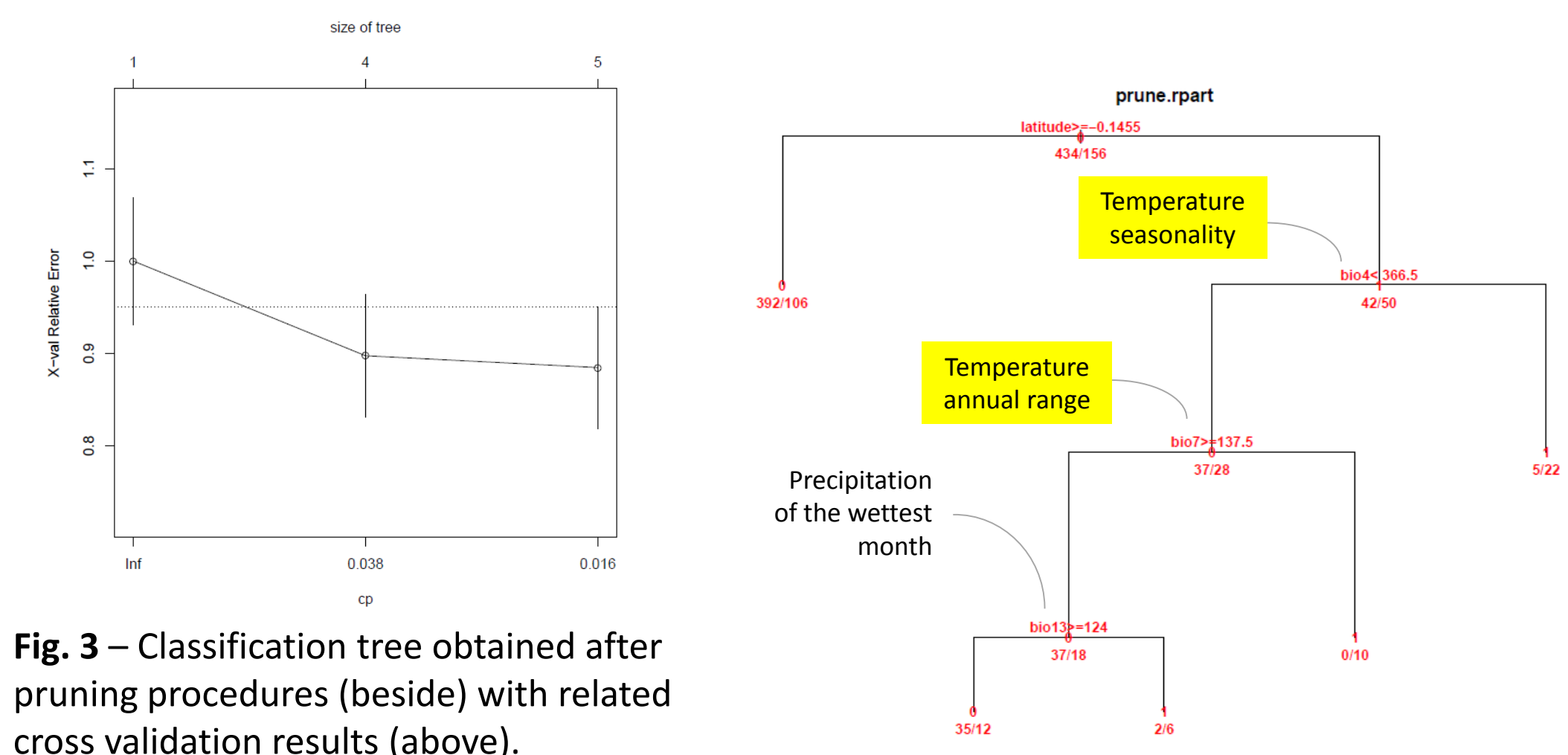


Fig. 3 – Classification tree obtained after pruning procedures (beside) with related cross validation results (above).

Predictor	Intercept	2.5	97.5	StdErr	Coefficient	2.5	97.5	StdErr	pvalue	Bonferroni
longitude	-1.1298	-1.3659	-0.8937	0.1205	-0.4205	-0.6444	-0.1965	0.1143	2.00E-04	0.0049
bio1	-1.1102	-1.3348	-0.8856	0.1146	-0.4693	-0.674	-0.2646	0.1045	0	1.00E-04
bio5	-1.1132	-1.3378	-0.8885	0.1146	-0.4697	-0.677	-0.2625	0.1057	0	2.00E-04
bio6	-1.111	-1.3359	-0.886	0.1148	-0.4739	-0.6758	-0.272	0.103	0	1.00E-04
bio8	-1.1109	-1.3373	-0.8844	0.1156	-0.4624	-0.6673	-0.2574	0.1046	0	2.00E-04
bio9	-1.1108	-1.3348	-0.8868	0.1143	-0.471	-0.6747	-0.2674	0.1039	0	1.00E-04
bio10	-1.1121	-1.3369	-0.8873	0.1147	-0.4666	-0.6732	-0.26	0.1054	0	2.00E-04
bio11	-1.1088	-1.3338	-0.8838	0.1148	-0.466	-0.6715	-0.2604	0.1049	0	2.00E-04
bio19	-1.1163	-1.3432	-0.8893	0.1158	-0.4413	-0.6498	-0.2328	0.1064	0	7.00E-04

Tab. 1 – Variables significantly associated with *T. parva* presence after Bonferroni correction ($P \leq 0.05$). In yellow, the variables also revealed by the classification tree in Fig. 3A; the gray rectangle emphasizes the estimated regression coefficients.

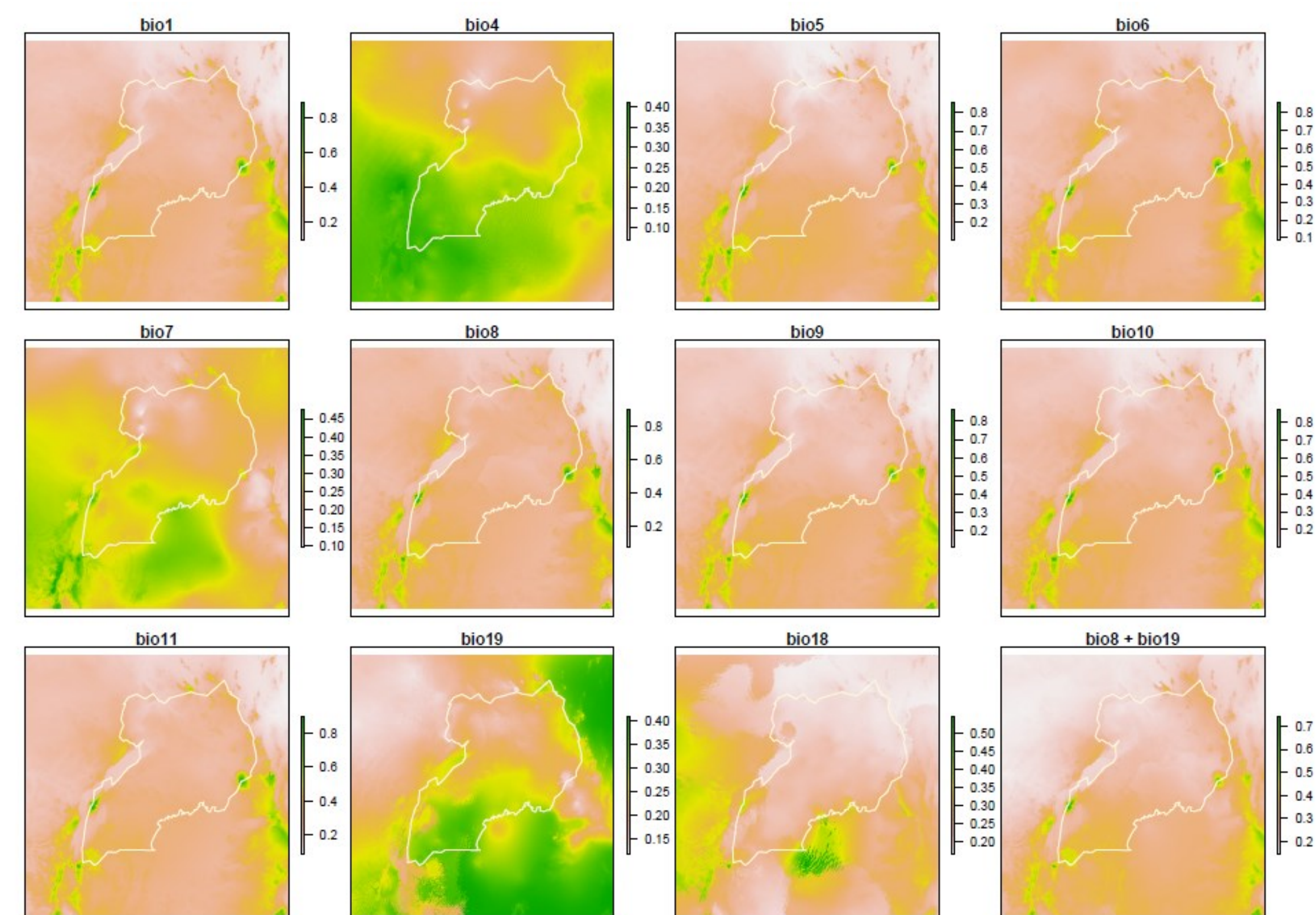


Fig. 4 – Probability of *T. parva* presence as modeled by the univariate and bivariate regressions. Each predictor is reported at the top of the maps. Probability is quantified by the bar beside the each map, and by the related color range. The bio8 + bio19 (Precipitation of Coldest Quarter) regression appears to be the best model on the basis of its AIC value (660.1). Map resolution is 30 arc-seconds ($\sim 1 \text{ Km}^2$).

Ongoing analyses and next steps

- Follow up with *T. parva* niche characterization by testing other Ecological Niche Modeling techniques.
- Study *R. appendiculatus* ecological niche, overlap its modeled distribution with the one of *T. parva* and try to elucidate the parasite-specific environmental requirements.
- Search for genomic regions conferring resistance to East Coast Fever in bovine host populations, by using a landscape genomics approach.

References

Bates D., Maechler M., Bolker B., & Walker S (2014) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. <http://CRAN.R-project.org/package=lme4>.
 Cumming G. S. (2002) Comparing Climate and Vegetation as Limiting Factors for Species Ranges of African Ticks. *Ecology*, **83**, 255–268.
 Dormann C. F., Elith J., Bacher S., Buchmann C., & *et al.* (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27–46.

Norval R.A.I., Perry B.D., & Young A.S. (1992) *The epidemiology of theileriosis in Africa*. ILRI (aka ILCA and ILRAD).
 Peterson A. P. & Jorge Soberón J. (2008) Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. *Natureza & Conservação*, **10**, 102–107.
 Ripley B. (2014) tree: Classification and regression trees. R package version 1.0-35. <http://CRAN.R-project.org/package=tree>.
 Therneau T., Atkinson B., & Ripley B. (2014) rpart: Recursive Partitioning and Regression Trees. R package version 4.1-8. <http://CRAN.R-project.org/package=rpart>.