10.1071/WR13028_AC

Supplementary Material: *Wildlife Research*, 2013, 40(7), 578-587.

Spatial prediction of brushtail possum (*Trichosurus vulpecula*) distribution using a combination of remotely sensed and field-observed environmental data

Thibaud Porphyre[A,D], *Joanna McKenzie*[A], *Andrea E. Byrom*[B], *Graham Nugent*[B,E], *James Shepherd*[C] *and Ivor Yockney*[B]

[A]EpiCentre, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand.

[B]Landcare Research, PO Box 69, Lincoln 8152, New Zealand.

[C]Landcare Research, Private Bag 11052, Palmerston North 4442, New Zealand.

[D]Present address: Epidemiology Group, Centre for Immunity, Infection and Evolution, University of Edinburgh, Ashworth Laboratories, Edinburgh, United Kingdom.

[E]Corresponding author. Email: nugentg@landcareresearch.co.nz

Remotely-sensed environmental databases used in model development

EcoSat database *EcoSat* (www.landcareresearch.co.nz/resources/maps-satellites/ecosat/):

Land cover variables were derived from EcoSat, 15 m resolution Landsat ETM+ satellite imagery that was processed to standardise spectral reflectance by correcting interferences caused by atmosphere, sun position, slope and view direction (Dymond and Shepherd 2004). An unsupervised classification was applied to the six standardised spectral reflectance bands generating 30 distinguishable classes (see below and Fig. S1).

Land Environments of New Zealand database *LENZ* (www.landcareresearch.co.nz):

LENZ is a classification of New Zealand's landscapes using a comprehensive set of climate variables, landform and soil variables that drive geographic variation in biological patterns (Leathwick *et al*. 2002). The most detailed level of classifications, containing 500 environments, was used in this study, of which 20 classes were present in the study area. The seven individual climate layers comprised within LENZ (i.e. mean annual temperature (°C), mean minimum temperature of the coldest month (°C), mean annual solar radiation ($MJ.m^{-2}d^{-1}$), winter solar radiation ($MJ.m^{-2}d^{-1}$), October vapour pressure deficit (KPa), annual water deficit (mm), and monthly water balance ratio were also considered in these analyses.

New Zealand Digital Elevation Model database *DEM* (http://geographx.co.nz/):

The DEM was used to derive mean altitude above sea level for each line. Mean height above valley floor (HVF) of each line was also generated from the DEM using an erosion terrains data set (Landcare Research Ltd, Lincoln, New Zealand) to identify valley floors, as this variable was thought to most closely reflect the local environmental conditions that influence the distribution of possums.

Processes followed for establishment of EcoSat classes used in the models

As no ground validation had been conducted, EcoSat classes that were recorded over the entire South Island were cross-tabulated with those recorded in the Land Cover Database, a land cover data set at 100 m resolution derived from satellite imagery that had been extensively validated in the field (http://www.mfe.govt.nz/issues/land/land-cover-dbase/). Cluster analysis was applied to the resulting matrix using the Bray-Curtis method to identify groups of EcoSat classes of similar LCDB-derived habitat composition (Leathwick *et al*., 2002). We determined the most appropriate clustering method for data by optimising the degree to which each hierarchy fitted the observed matrix using the cophenetic correlation coefficient (CCC; Romesburg 1984, Tan *et al*. 2005; i.e. the Pearson's product-moment correlation between the tested hierarchies) . The degree of fit of the actual data was assessed for seven different clustering methods: the group average method (UPGMA), median linkage method, centroid method, Ward's minimum variance clustering method, McQuitty's method, complete linkage method, and single linkage method (Romesburg 1984; Tan *et al*. 2005). The UPGMA method resulted in the most accurate hierarchical clustering tree (CCC = 0.85), producing seven communities of EcoSat classes that had similar LCDB2 composition plus six individual classes, giving a total of 13 habitat clusters in the EcoSat data (Fig. S1).

References

Dymond J, Shepherd J (2004) The spatial distribution of indigenous forest and its composition in the Wellington region, New Zealand, from ETM+ satellite imagery. Remote Sensing of Environment **90**, 116-125.

Leathwick J, Morgan F, Wilson G, Rutledge D, McLeod M, Johnston K (2002) Land environments of New Zealand: a technical guide. New Zealand Ministry for the Environment online publication, pp 1 -237. http://www.landcareresearch.co.nz/__data/assets/pdf_file/0020/21773/LENZ_Technical_Guide.pdf

Romesburg H (1984) Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont, CA.

Tan P-N, Steinbach M, Kumar V (2005) Introduction to Data Mining 1st edn. Addison Wesley Publishing Company, New York.

Figure S1. Dendrogram showing clusters of 30 EcoSat habitat classes with similar LCDB2 (national land cover database) composition. Clusters involving multiple EcoSat classes with a similar composition are indicated by red boxes. Each cluster plus individual EcoSat classes were labelled in alphabetical order (bottom letters).

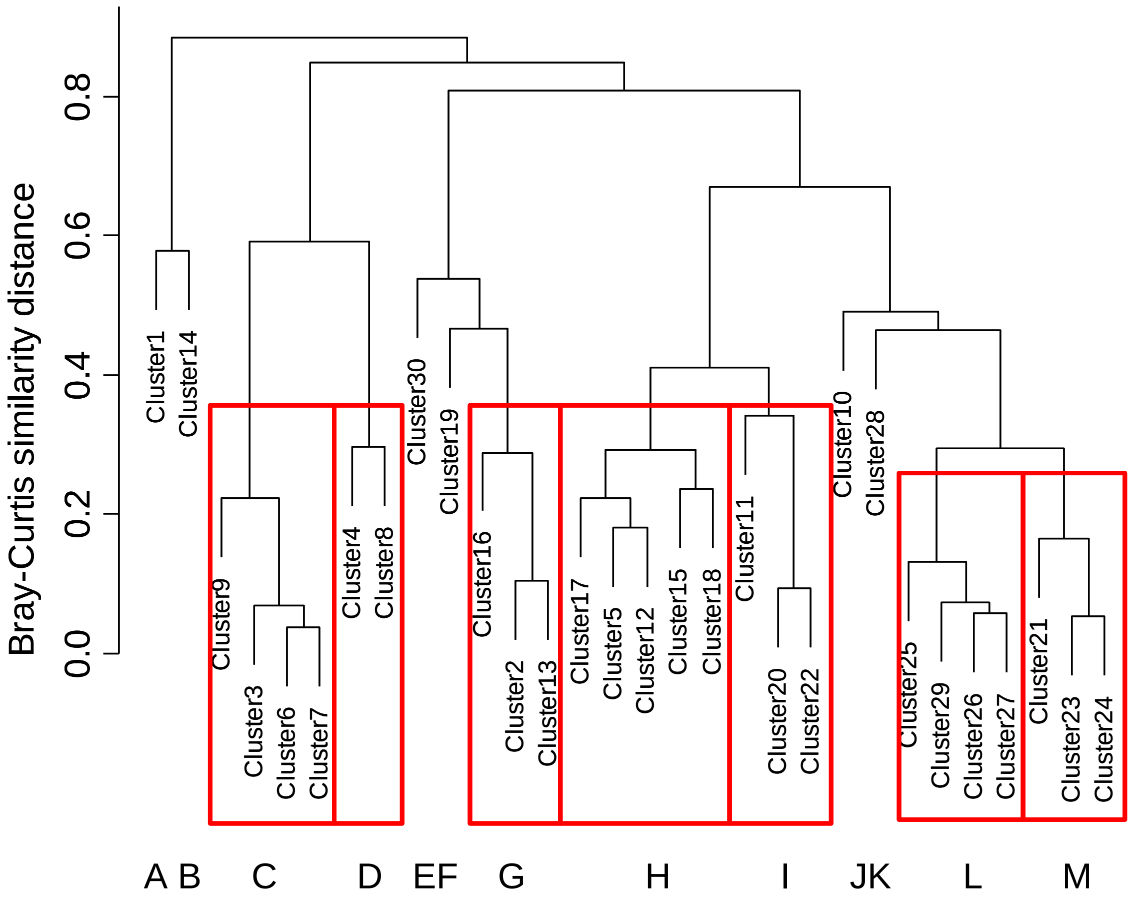

Figure S2. Normal quantile-quantile (Q-Q) plots of deviance residuals for two multivariate negative binomial regression models (DIGITAL (a) and DIGITALFIELD (b)) used to predict relative abundance of possums on Molesworth Station in the northern South Island high country of New Zealand. If the residual values are reasonably well approximated by a normal distribution, all points would fall along the solid straight line, while points outside the 95% confidence envelope (dashed lines) mark deviations from this hypothesis.

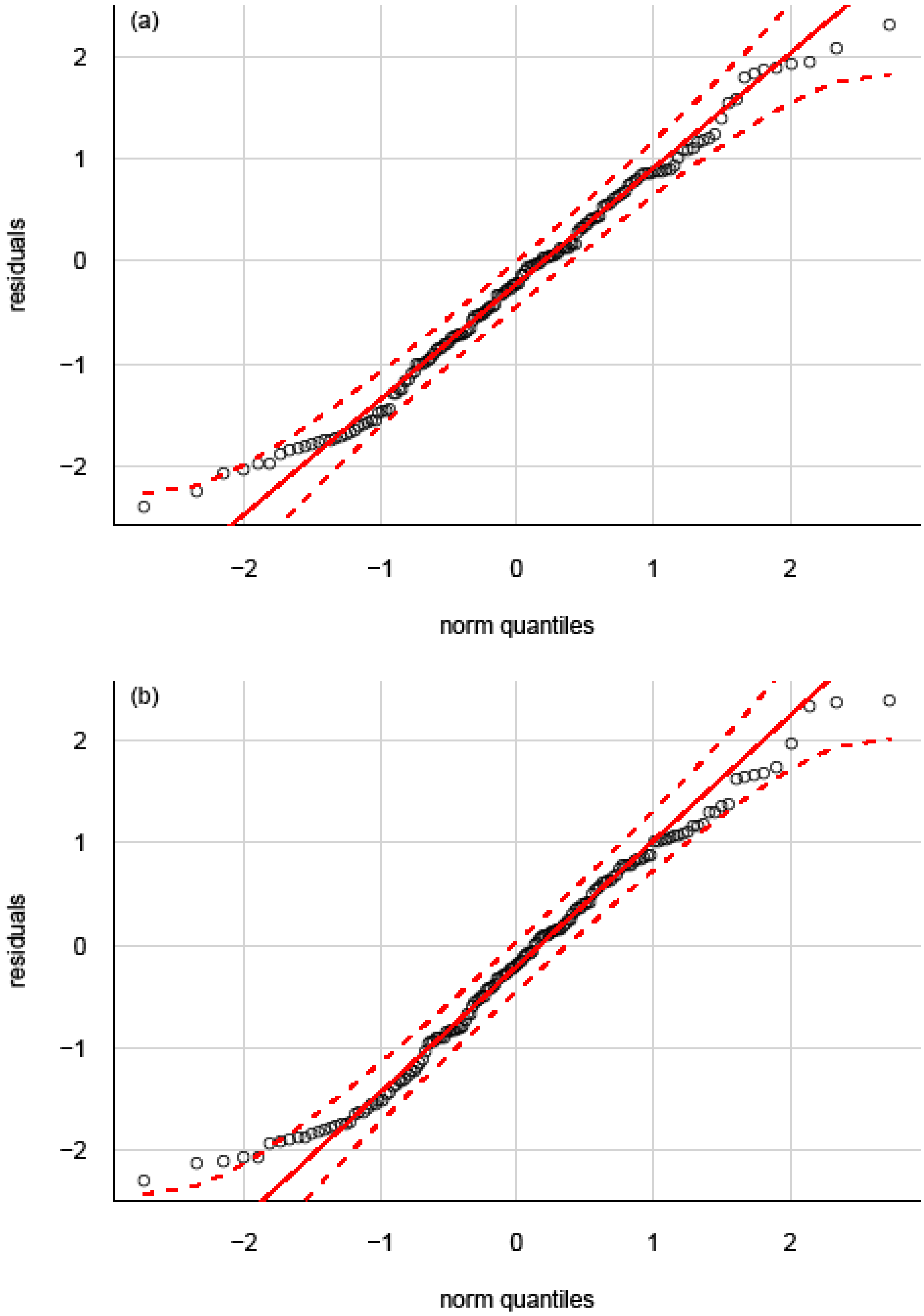

Figure S3. Spatial map of the Molesworth Station study site depicting error around the possum relative abundance predictions for illustrative purposes (error represented here by mapped coefficients of variation). The bar scale at the right of the figure represents the range of coefficients of variation over the mapped area.