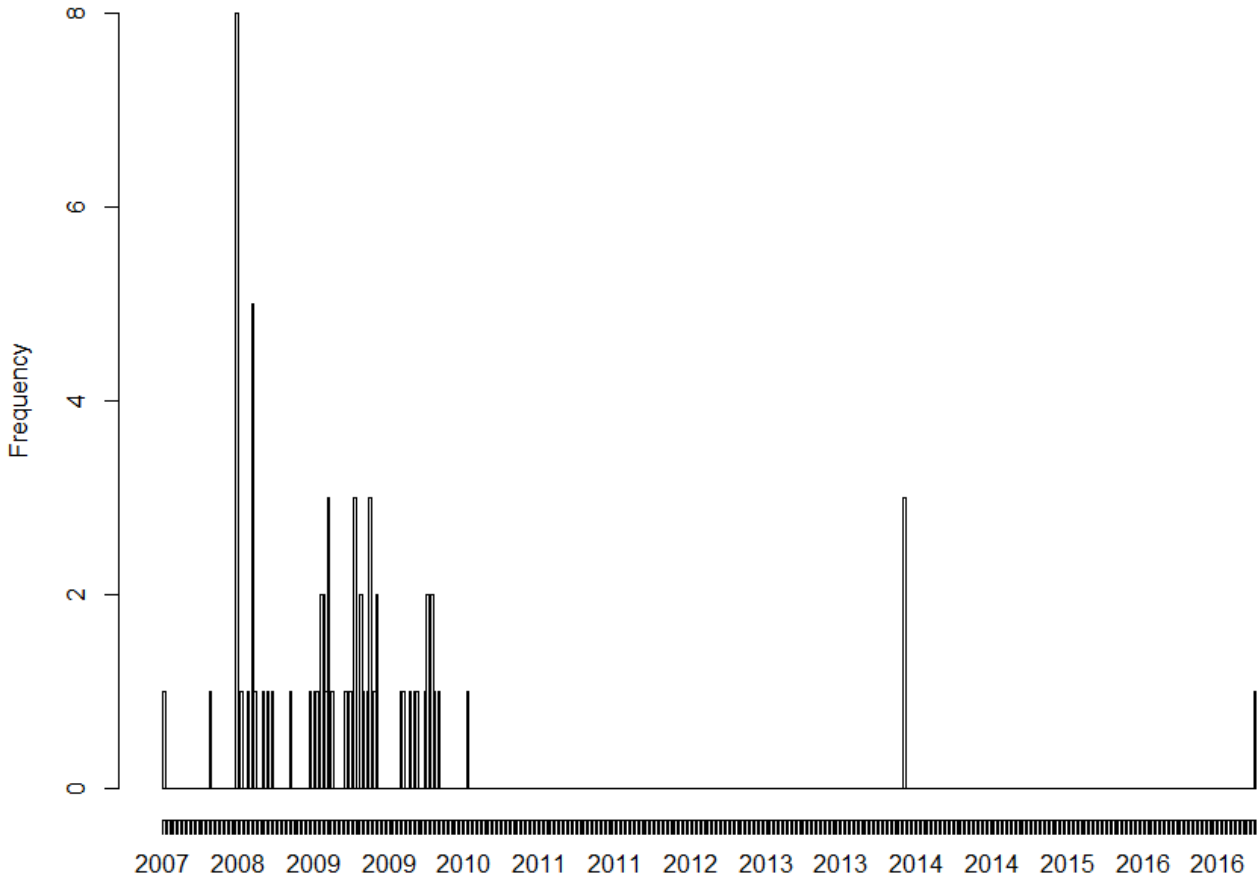
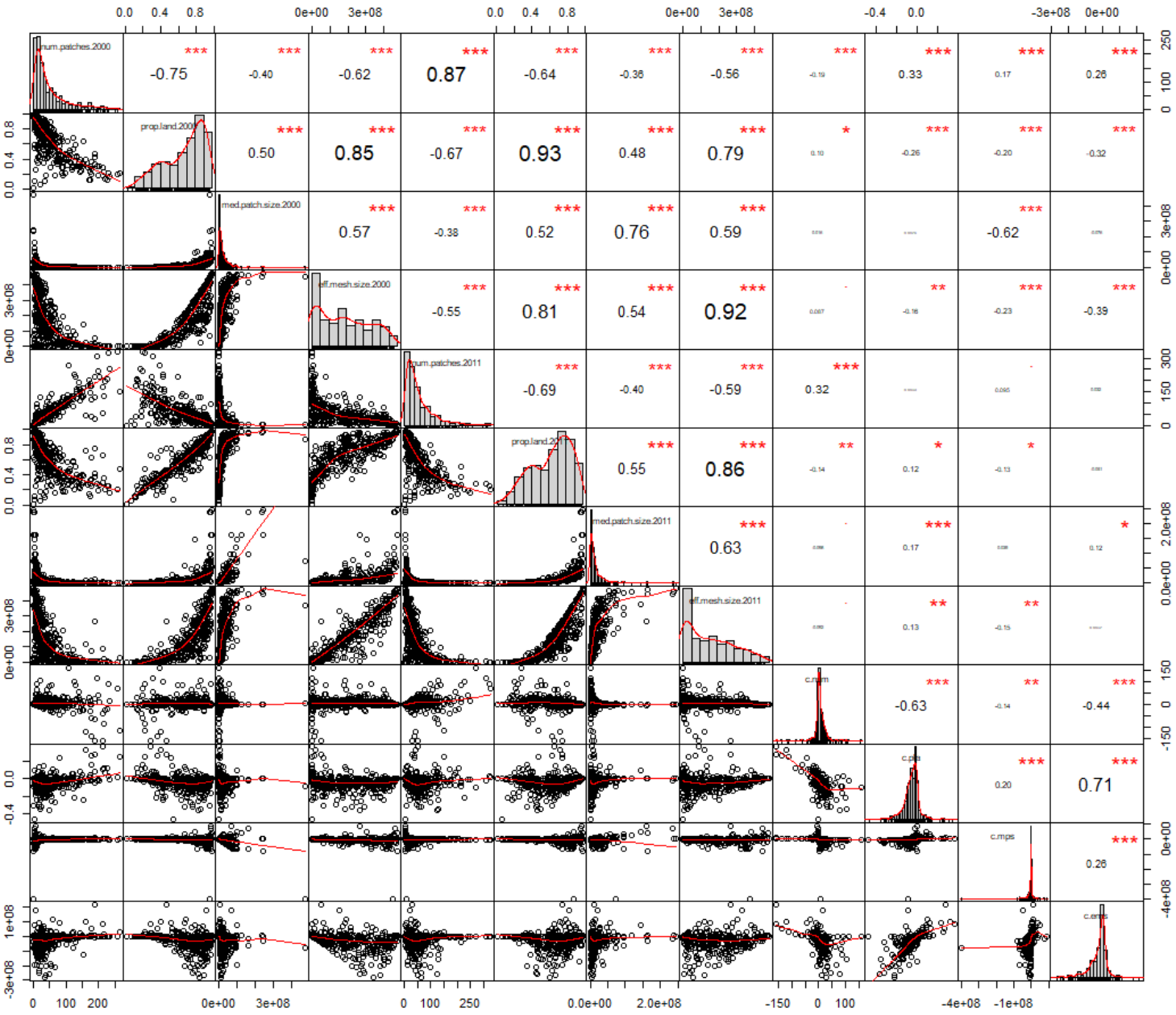
Supplementary Information

Synergic effects of grassland fragmentation and temperature on bovine rabies emergence

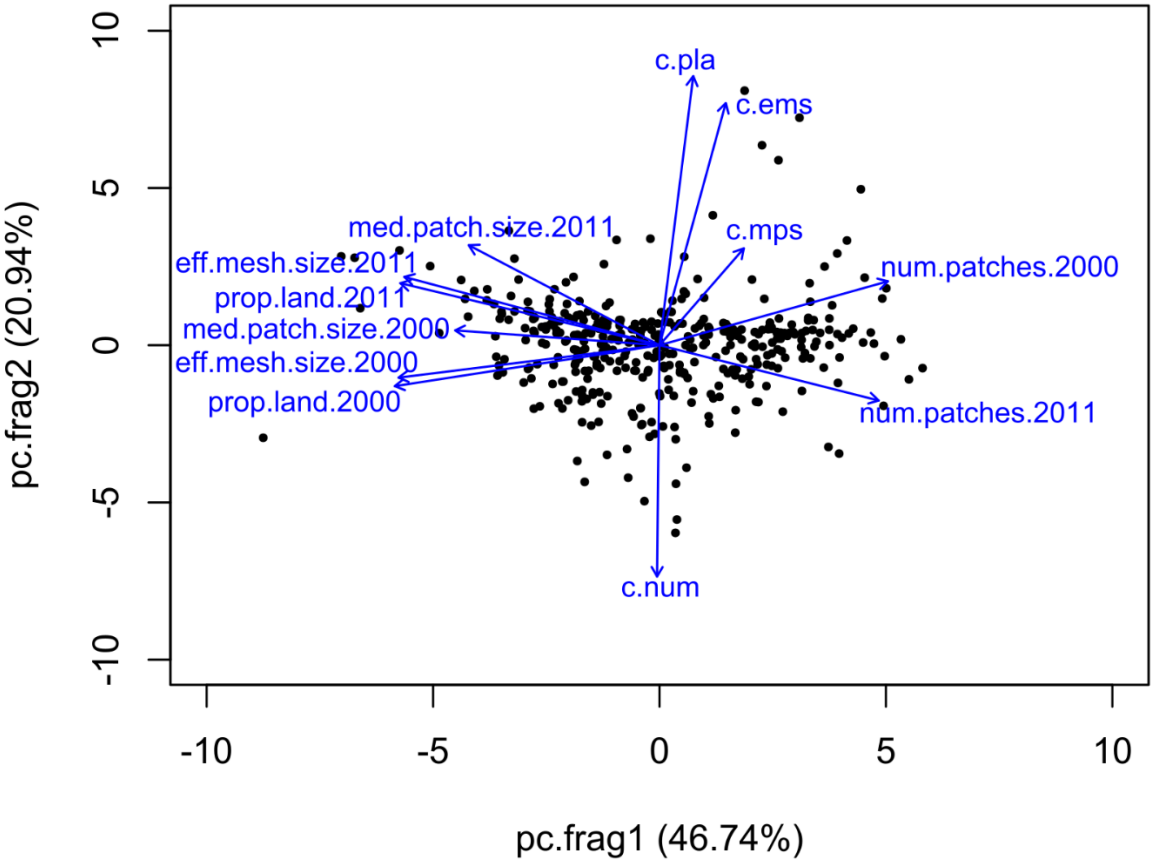
Supplementary Figure S1. Temporal distribution of rabies outbreaks as number of new quarantined ranches per week from 2007 to 2017.



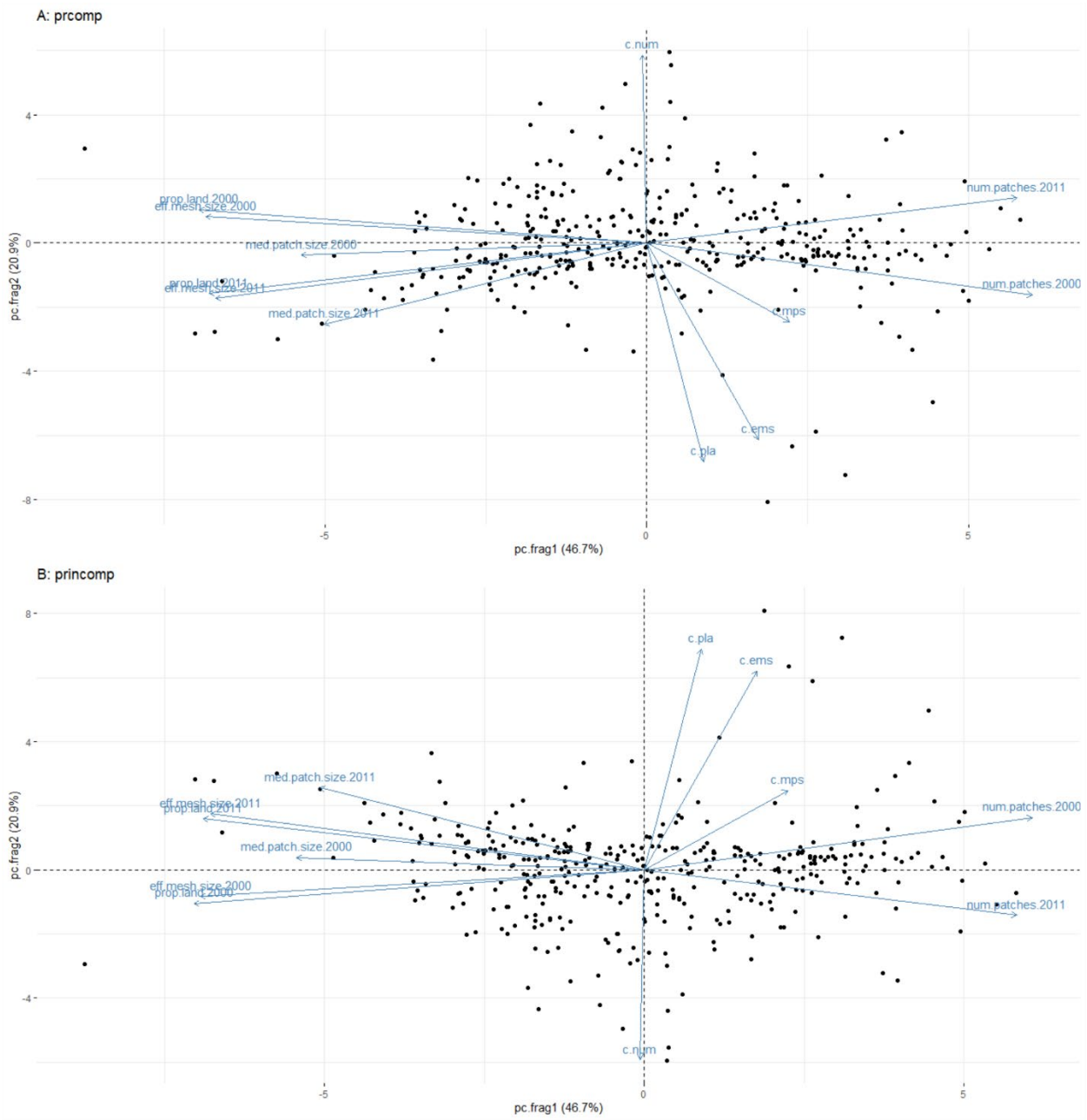
Supplementary Figure S2. Correlation matrix of grassland fragmentation metrics. Collinearity is observed for several pairs of variables, showing that variables provide similar information about the fragmentation process



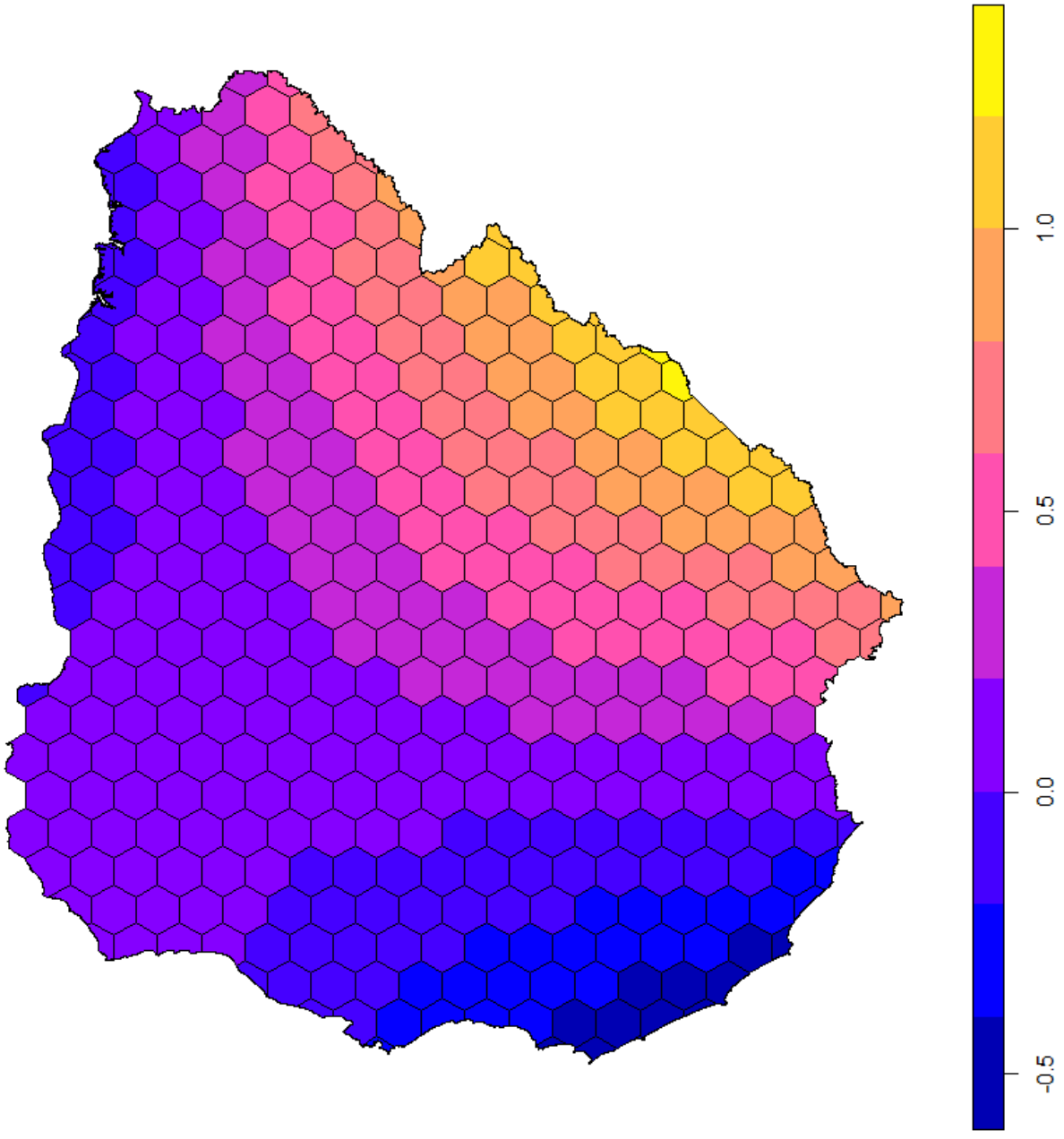
Supplementary Figure S3. Biplot for the Principal Component Analysis on the grassland fragmentation metrics. The first principal component (*pc.frag1*) explains the 46.74% of the total variance and is related to fragmentation status. *pc.frag1* increases with the increase in number of patches or with the decrease in proportion of landscape covered by livestock areas, the mean patch size and the effective mesh size. The second principal component (*pc.frag2*) explains the 20.94% of the total variance and is associated with recent change in fragmentation. Negative values in the second principal component are indicative of increased fragmentation from 2000 to 2011. Variables included in the analysis: *num.patches.2000*, *num.patches.2011* (number of grassland patches per cell in 2000 and 2011), *prop.land.2000*, *prop.land.2011* (proportion of cell covered by grasslands in 2000 and 2011), *med.patch.size2000*, med.patch.size2011 (mean size of grassland patches per cell in 2000 and 2011), *eff.mesh.size2000*, *eff.mesh.size2011* (effective mesh size of grassland areas in 2000 and 2011), *c.num* (change in number of grassland patches from 2000 to 2011), *c.pla* (change in proportion of cell covered by grasslands patches from 2000 to 2011), *c.mps* (change in mean size of grassland patches per cell from 2000 to 2011), *c.ems* (change in effective mesh size of grassland patches per cell from 2000 to 2011). See text for calculation of grassland fragmentation metrics.



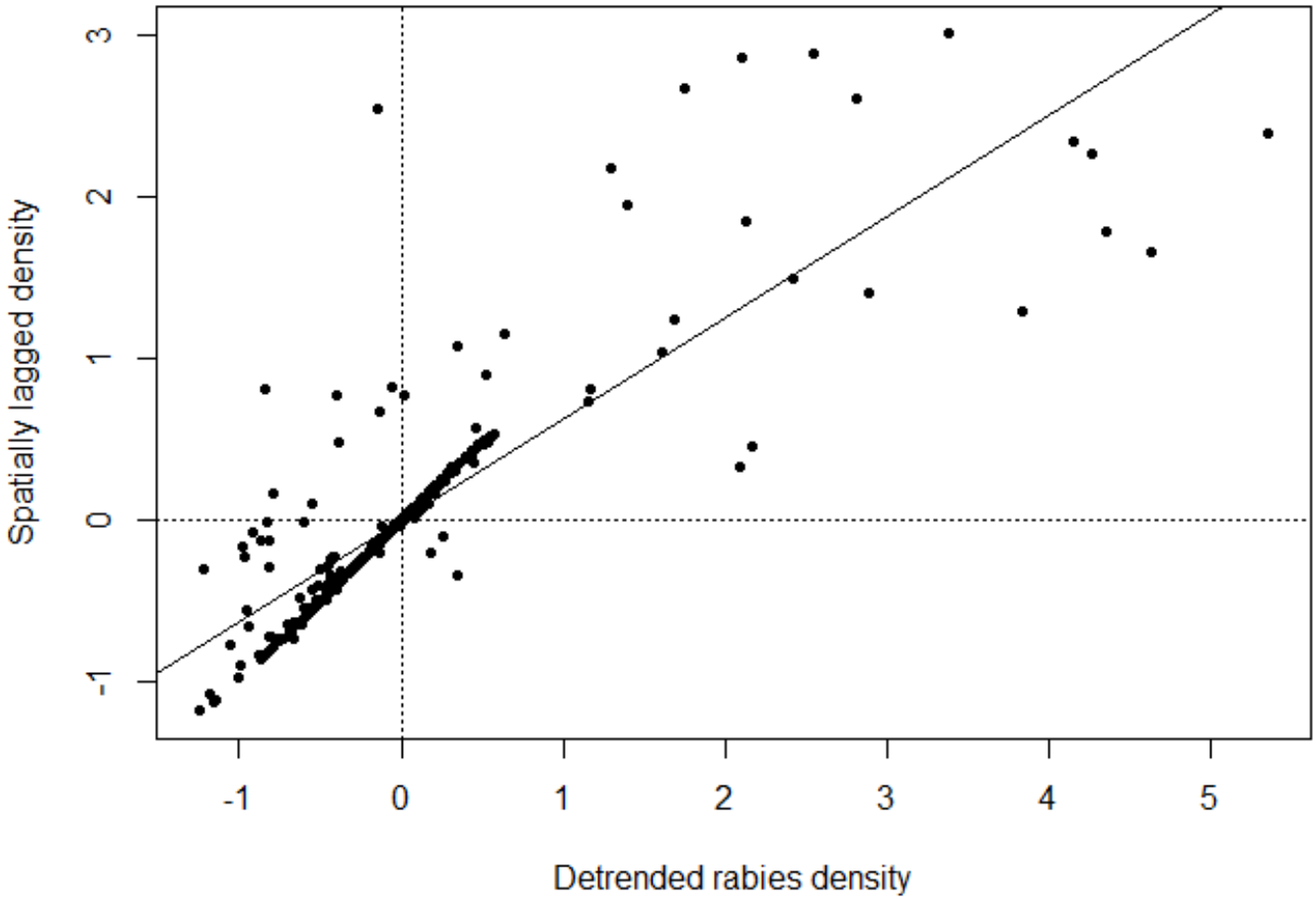
Supplementary Figure S4. Comparison of the results from the Principal Component Analyses using spectral decomposition (function prcomp, **panel A**) and singular decomposition (function princomp, **panel B**). Both methods gave similar results, with only inverted loadings on the second principal component.



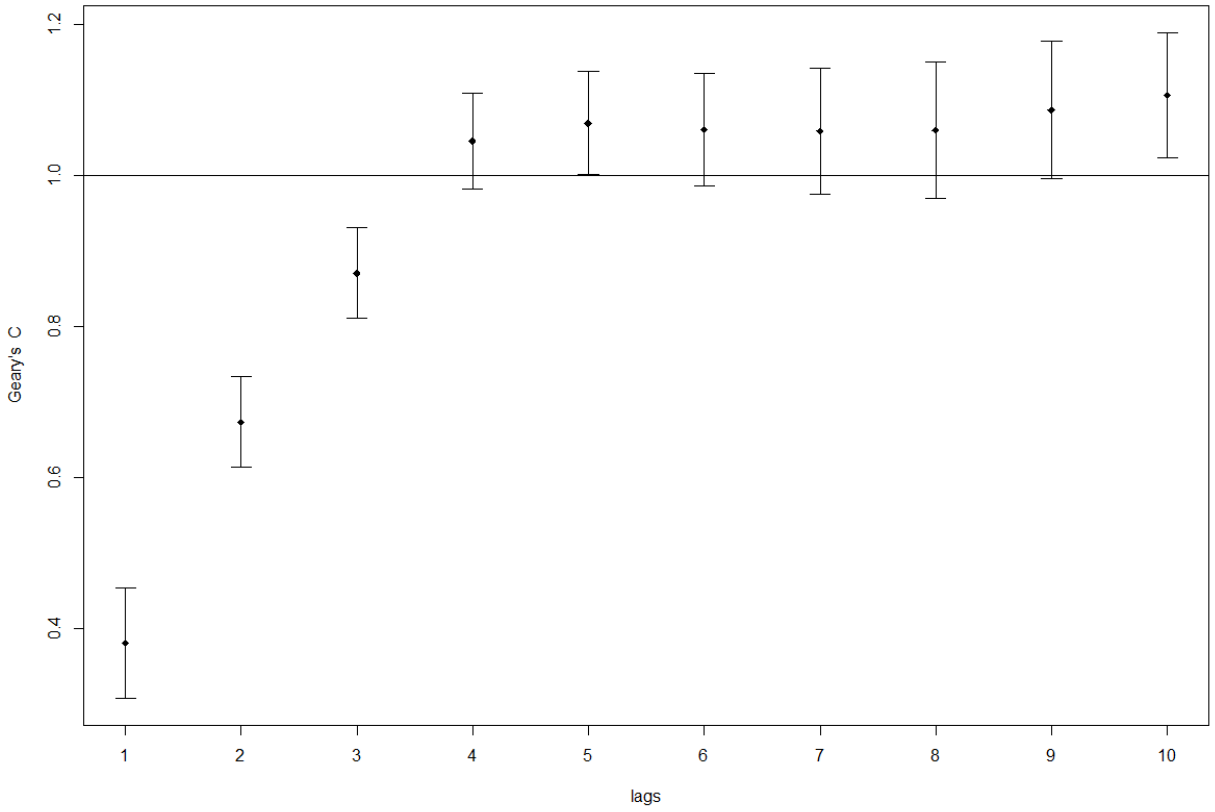
Supplementary Figure S5. Distribution of fitted values from the linear regression showing a northeast-to-southwest trend on the density of ranches with rabies cases. There are higher expected densities in the Northeastern region of the country and decreasing values in the northeast-to-southwest direction. See also Figure 3 A-B in the main text for a comparison of the predicted values with the observed density of quarantined ranches.



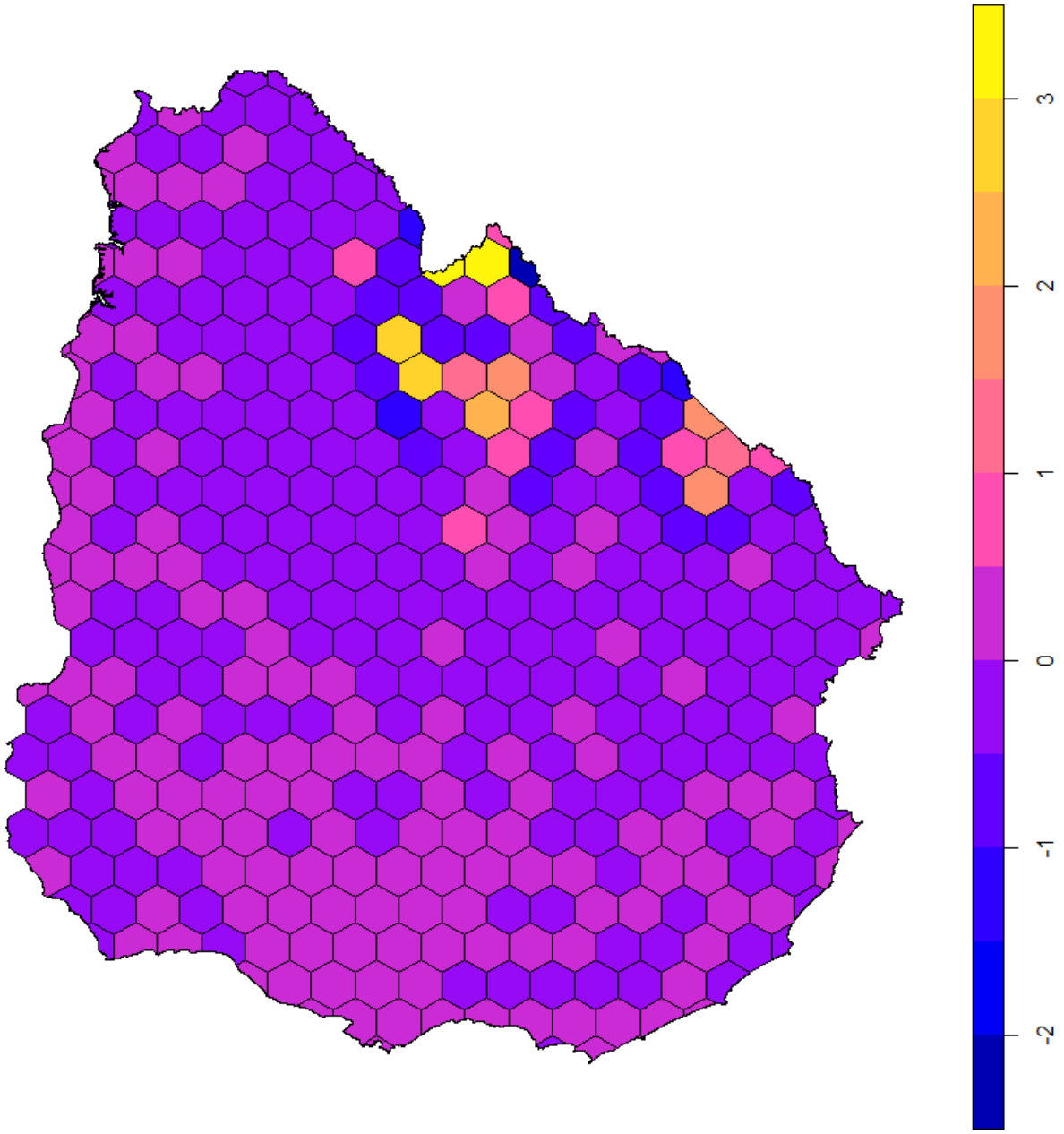
Supplementary Figure S6. Moran plot for the residuals from the linear regression analysis. Even after taking out the first order effect, the detrended residuals still show a positive spatial autocorrelation (C=0.38, p-value<0.01), showing that there is an underlying spatial structure of the rabies outbreaks not explained by the invasion-wave-like effect.



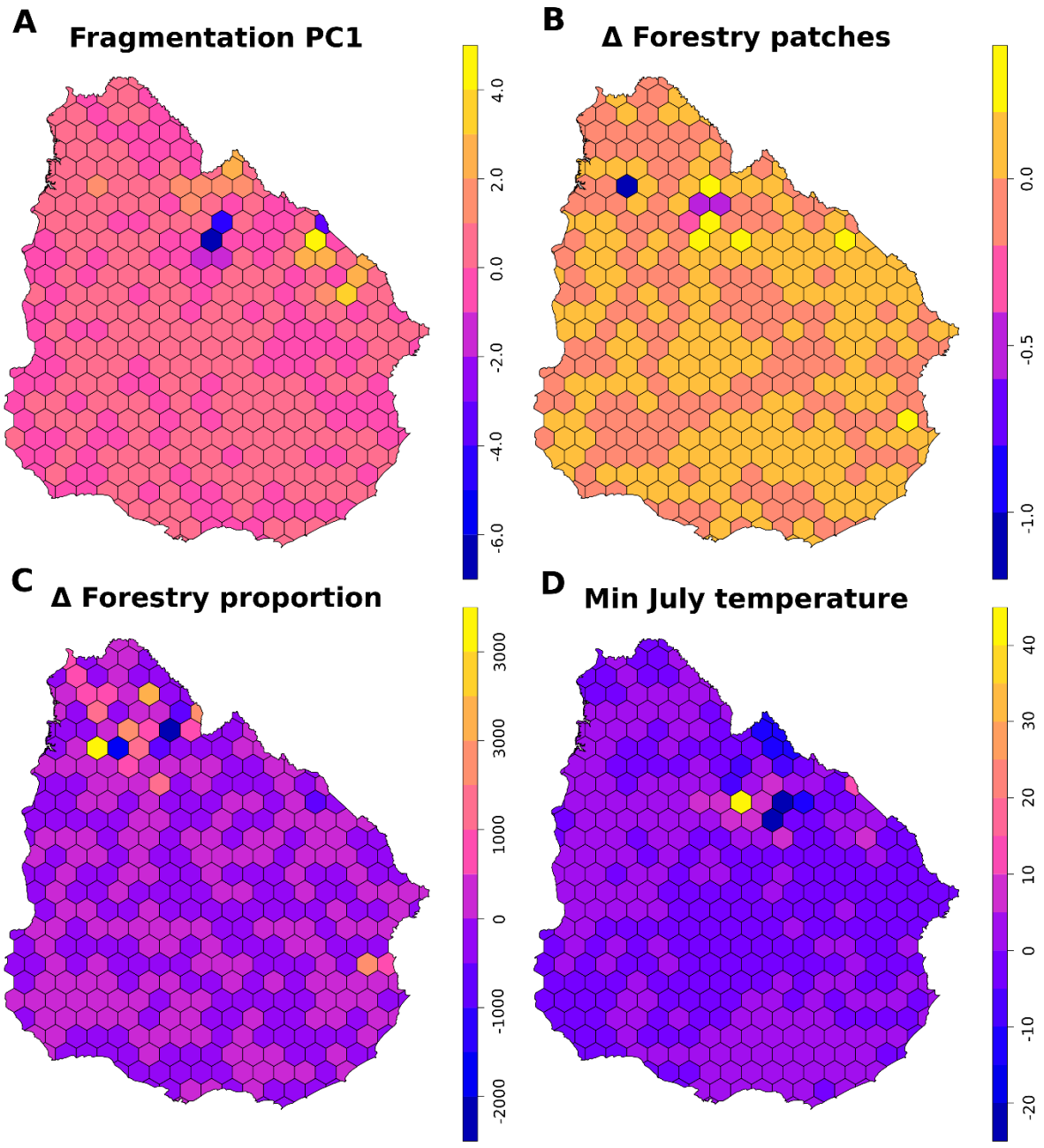
Supplementary Figure S7. Correlogram for the residuals from the linear model. A positive significant autocorrelation (Geary's C value lower than 1) is observed only for the first three lags, whereas for the rest of the lags, no significant autocorrelations are observed. The spatial structure of density of outbreaks is only observed up to the third order of neighboring in the regular hexagonal-cell grid.



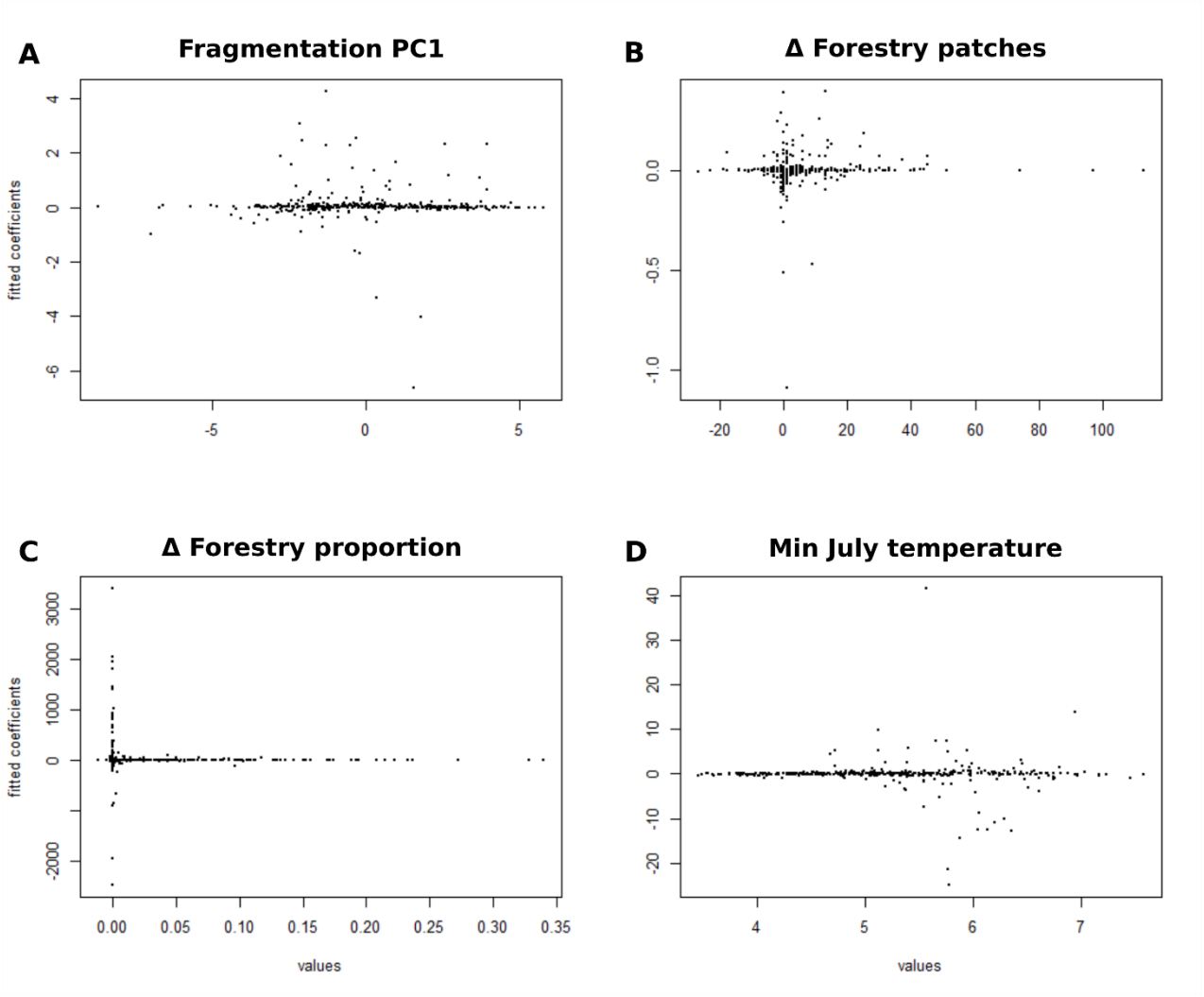
Supplementary Figure S8. Distribution of the residuals from the final Spatial Autoregressive (SAR) Model (*detrended_residuals ~ c.fnp * temp_jul*). The autocorrelation analysis for the residuals, using Geary's C, shows no significant autocorrelation (Geary's C=0.96, p-value=0.15).



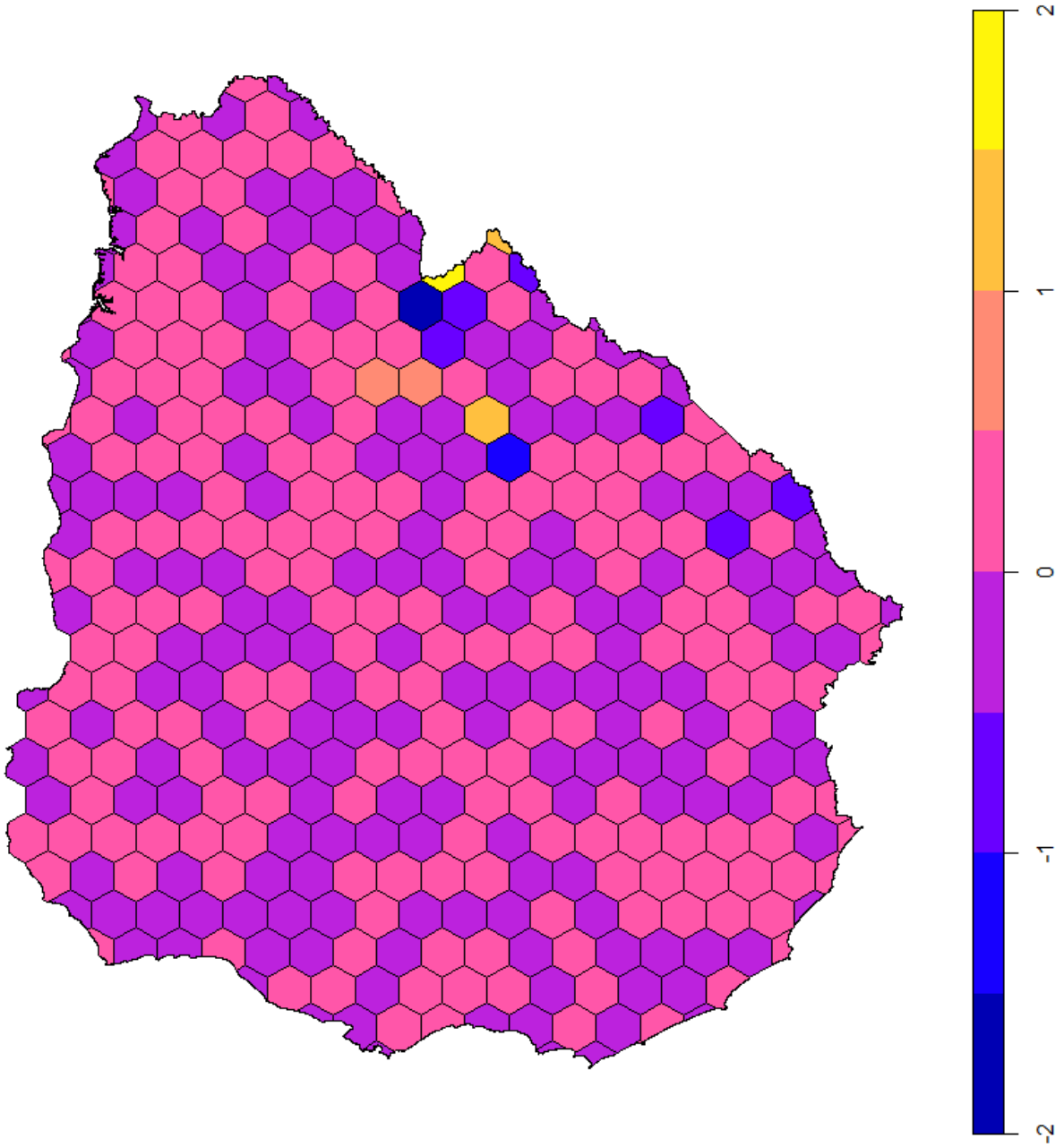
Supplementary Figure S9. Results from the Geographically Weighted Regression model: geographic distribution of fitted coefficients from the final model (*detrended_residuals ~ pc.frag1 + c.fnp + c.fpl + temp_jul*). **A.** Fragmentation PC1 (*pc.frag1*), positive values indicate higher fragmentation of livestock areas **B.** Δ forestry patches (*c.fnp*), change in number of exotic forestry patches between 2000 and 2011 **C.** Δ forestry proportion (c.fpl), change in proportion of landscape covered by exotic forestry between 2000 and 2011 **D.** Minimum Jul temperature (*temp_jul*), average minimum temperature of the coldest month.



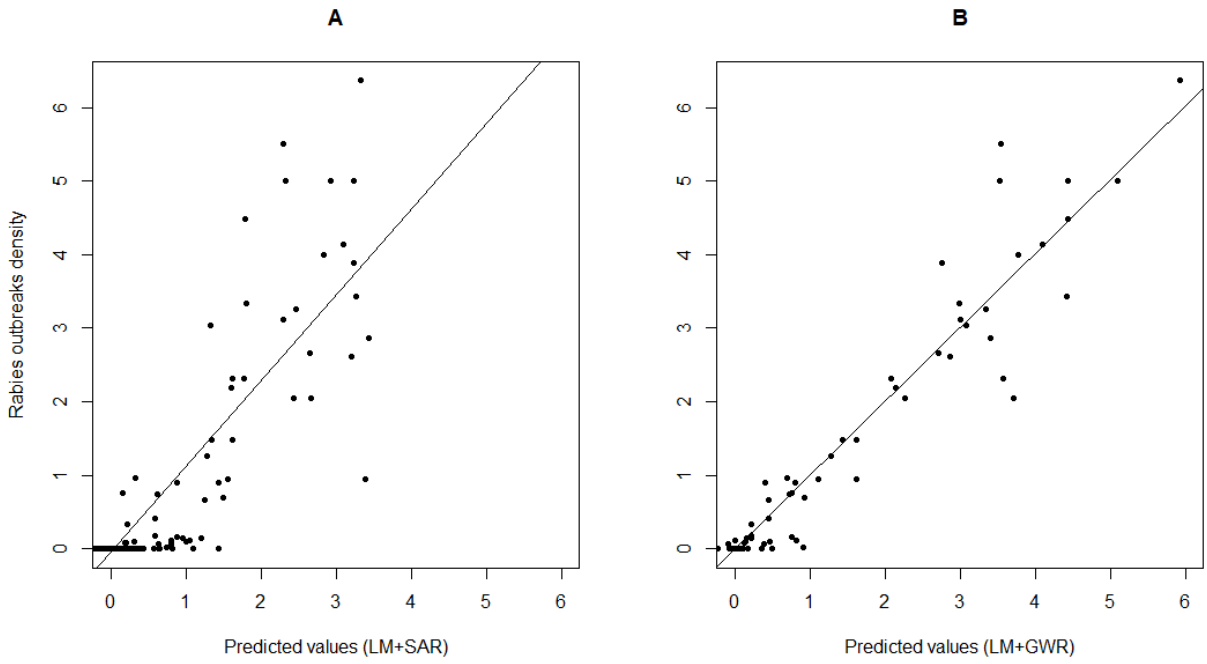
Supplementary Figure S10. Distribution of the fitted coefficient values for the final GWR model (*detrended_residuals ~ pc.frag1 + c.fnp + c.fpl + temp_jul*), in relation to the original values of each variable: **A.** Fragmentation PC1 (*pc.frag1*), positive values indicate higher fragmentation of livestock areas **B.** Δ forestry patches (*c.fnp*), change in number of exotic forestry patches between 2000 and 2011 **C.** Δ forestry proportion (c.fpl), change in proportion of landscape covered by exotic forestry between 2000 and 2011 **D.** Minimum Jul temperature (*temp_jul*), average minimum temperature of the coldest month.



Supplementary Figure S11. Distribution of the residuals from the final Geographically Weighted Regression (GWR): *detrended_residuals ~ pc.frag1 + c.fnp + c.fpl + temp_jul*. Residuals from the final model showed no significant spatial autocorrelation (Geary's C=1.13, p-value=0.99).



Supplementary Figure S12. Linear correlation of rabies outbreak density with predicted values by combination of linear model and Simultaneous Autoregressive model (A), and combination of linear model and Geographically weighted regression (B). The plotted lines in both cases represent the simple linear regression between the rabies outbreaks density and the predicted values. In both cases there is a positive and significative linear association (R^2=0.76, p<0.01 and R^2=0.94, p<0.01 respectively), showing good prediction power of both methods. Not only there is a strong correlation, but also in both examples, the intercept is not different from zero (p>0.05) and the regression coefficients for the predicted values are close to 1 (β=1.16 for SAR and β=1.00 for GWR), showing that there almost no bias in the prediction. In B, the GWR shows better accuracy and less dispersion of predicted values for each observed value. Non-stationary coefficients allow the model to perform better than SAR in predicting the zero values for rabies density.



Supplementary Table S1. Legend of the Land Cover Classification System used in Uruguay (LCCS-Uy, modified from Álvarez et al. 2015). The four classes considered for livestock areas (i.e. He, Ar, Pa and ANi, by their original codes in Spanish) and the class for forestry (PF) are highlighted. We selected the classes to include at the 17-classes level, as the 46-classes legend is not conserved across different versions of the LCCS-Uy products.

Main Groups	17-classes legend	46-classes legend
A11 - Cultivated and managed terrestrial areas	Irrigated crops >4-5 ha	Irrigated crops >4-5 ha
		Sugar cane
		Rice field >4-5 ha
		Sugar cane or rice > 4-5 ha
	Rainfed crops > 4-5 ha	Rainfed crops > 4-5 ha
	Small crops <4-5 ha	Rainfed crops < 4-5 ha
		Irrigated crops < 4-5 ha
	Forestry Plantations (PF)	Forestry venture > 5 ha
		Coastal implanted forest
		Eucalyptus forest > 5 ha
		Pine trees forest > 5 ha
		Shelter and shadow forest <5 ha
		Urban park
	Orchards	Citrus orchards
		Fruit orchards
A12 - Natural and semi-natural vegetation	**Natural Herbaceous (He)**	Natural grassland
		Sandplain herbaceous
		Natural grassland with sparse palm groves (1-15%)
		Natural grassland with rocks
	Shrublands (Ar)	Shrublands and natural grassland
	Native Forests	Native hillside forests
		Native riparian forest
		Native forests
		Native savannah forest
	Palm Groves (Pa)	Palm Groves
A24 - Natural and semi-natural aquatic or regularly flooded vegetation	**Natural Flooded Areas (ANi)**	Permanently flooded herbaceous coverage
		Temporarily flooded herbaceous coverage
B15 - Artificial surfaces and associated areas	Urban equipment	Airport
		Airfield
		Sport facilities
		Industrial facilities
		Port areas
	Urban areas	Urban areas
	Sparse Urban Areas	Sparse urban area and croplands
		Sparse urban area and natural grasslands
		Sparse urban area and forestry
B16 - Bare soil areas	Quarries, sand pits, open-pit mines	Quarries, sand pits, open-pit mines
	Bare soil areas	Sand beaches
		Sand dunes
		Consolidated rock
		Bare soil
B27 - Artificial ice, snow or water bodies	Artificial water bodies	Channels
		Lakes, reservoirs
B28 - Natural ice, snow or water bodies	Natural water bodies	Lagoons
		Water courses
		Humid and seasonally flooded soils

Supplementary Table S2. Variables of livestock areas fragmentation included in the Principal Component Analysis and loading scores for first two principal components (*pc.frag1* and *pc.frag2*, very small loadings are not printed, and substituted by a star). Number of grassland patches per cell (in 2000, 2011 and the difference 2011-2000: *num.patches.2000*, *num.patches.2011*, *c.num*), Proportion of grassland landscape (in 2000, 2011 and the difference 2011-2000: *prop.land.2000*, *prop.land.2011*, *c.pla*), Mean grassland patch size (in 2000, 2011 and the difference 2011-2000: *med.patch.size.2000*, *med.patch.size.2011*, *c.mps*), and Effective grassland mesh size (in 2000, 2011 and the difference 2011-2000: *eff.mesh.size.2000*, *eff.mesh.size.2011*, *c.ems*)

		Loading scores	
Variable description	**Variable name**	***pc.frag1***	***pc.frag2***
Number of grassland patches per cell 2000	*num.patches.2000*	0.336	0.136
Number of grassland patches per cell 2011	*num.patches.2011*	0.323	-0.117
Change in number of grassland patches per cell (2011-2000)	*c.num*	*	-0.490
Proportion of grassland landscape 2000	*prop.land.2000*	-0.390	*
Proportion of grassland landscape 2011	*prop.land.2011*	-0.382	0.132
Change in the proportion of grassland landscape (2011-2000)	*c.pla*	*	0.570
Mean grassland patch size 2000	*med.patch.size.2000*	-0.301	*
Mean grassland patch size 2011	*med.patch.size.2011*	-0.281	-0.281
Change in mean grassland patch size (2011-2000)	*c.mps*	0.124	0.205
Effective grassland mesh size 2000	*eff.mesh.size.2000*	-0.384	*
Effective grassland mesh size 2011	*eff.mesh.size.2011*	-0.376	0.145
Change in effective grassland mesh size (2011-2000)	*c.ems*	*	0.513

Supplementary Table S3. Results from the linear model using the latitude (Y), longitude (X) and the interaction between both (Y * X). Estimated coefficients for the two variables and the double interaction are presented, along with the p-values (starred p-values are significant at 0.05 level).

Variables	**Estimate**	**p-value**
Intercept	5.16e+01	
Y	-8.21e-06	* <0.01
X	-1.14e-04	* <0.01
*Y * X*	1.82e-11	* <0.01

Supplementary Table S4. Descriptive statistics for the values from the coefficients in the last GWR model (*detrended_residuals ~ pc.frag1 + c.fnp + c.fpl + temp_jul*). The spatial structure of the coefficients is presented in the values of Geary's C autocorrelation index and the correspondent p-values (starred p-values are significant at 0.05 level).

Variables	Min	Median	Max	Geary's C	p-value
pc.frag1	-6.64	0.01	4.26	0.75	*<0.01
c.fnp	-1.09	0.00	0.40	0.99	0.35
c.fpl	-2473.10	0.00	3400.60	1.02	0.68
temp_jul	-0.25	-0.01	41.56	0.76	*<0.01

Supplementary table S5. Results from the linear models to assess concordance among observed rabies outbreaks density and combined predictions from linear model (trend) and either Simultaneous Autoregressive model (SAR) or Geographically Weighted Regression (GWR). Estimated coefficients for the predicted values in each model combination are presented, along with the p-values (starred p-values are significant at 0.05 level). The second model show better fitting (R^2=0.94), concordance (coefficient estimate=1) and no bias (intercept estimate = 0). Figure S9 shows graphically the perform of both analyses.

Variables	**Estimate**	**p-value**	**R^2**
Model 1: *rabies density ~ (LM + SAR predicted values)*			0.76
Intercept	-0.04	0.09	
LM + SAR predicted values	1.16	*<0.01	
Model 2: *rabies density ~ (LM + GWR predicted values)*			0.94
Intercept	0.00	0.66	
LM + GWR predicted values	1.00	*<0.01	