remote sensing

MDPI

Article

On the Synergistic Use of Optical and SAR Time-Series Satellite Data for Small Mammal Disease Host Mapping

Christopher Marston [1,2,*] and Patrick Giraudoux [3]

[1] Department of Geography, Edge Hill University, St. Helens Road, Ormskirk, Lancashire L39 4QP, United Kingdom
[2] Centre for Ecology and Hydrology, Library Avenue, Lancaster Environment Centre, Bailrigg, Lancaster LA1 4AP, United Kingdom.
[3] Department of Chrono-environment, University of Bourgogne Franche-Comte/CNRS, La Bouloie 25030 Besançon Cedex, France.; patrick.giraudoux@univ-fcomte.fr
* Correspondence: marstonc@edgehill.ac.uk; Tel.: +44-1695-654-396

Received: 29 November 2018; Accepted: 21 December 2018; Published: 27 December 2018

Abstract: (1) Background: *Echinococcus multilocularis* (Em), a highly pathogenic parasitic tapeworm, is responsible for a significant burden of human disease. In this study, optical and time-series Synthetic Aperture Radar (SAR) data is used synergistically to model key land cover characteristics driving the spatial distributions of two small mammal intermediate host species, *Ellobius tancrei* and *Microtus gregalis*, which facilitate Em transmission in a highly endemic area of Kyrgyzstan. (2) Methods: A series of land cover maps are derived from (a) single-date Landsat Operational Land Imager (OLI) imagery, (b) time-series Sentinel-1 SAR data, and (c) Landsat OLI and time-series Sentinel-1 SAR data in combination. Small mammal distributions are analyzed in relation to the surrounding land cover class coverage using random forests, before being applied predictively over broader areas. A comparison of models derived from the three land cover maps are made, assessing their potential for use in cloud-prone areas. (3) Results: Classification accuracies demonstrated the combined OLI-SAR classification to be of highest accuracy, with the single-date OLI and time-series SAR derived classifications of equivalent quality. Random forest analysis identified statistically significant positive relationships between *E. tancrei* density and agricultural land, and between *M. gregalis* density and water and bushes. Predictive application of random forest models identified hotspots of high relative density of *E. tancrei* and *M. gregalis* across the broader study area. (4) Conclusions: This offers valuable information to improve the targeting of limited-resource disease control activities to disrupt disease transmission in this area. Time-series SAR derived land cover maps are shown to be of equivalent quality to those generated from single-date optical imagery, which enables application of these methods in cloud-affected areas where, previously, this was not possible due to the sparsity of cloud-free optical imagery.

Keywords: *Echinococcus multilocularis*; *Ellobius tancrei;* land cover; *Microtus gregalis*; random forests; SAR; Sentinel; spatial epidemiology; time-series

Table S1. Confusion matrix for the SAR and OLI combined classification. AG = agriculture, AP = alpine grassland, BR = bare, BU = built up, BS = bushes, DG = dry grassland, ST = steppe, WA = water.

Classified Data	**Reference Data**								
	BR	**BU**	**WA**	**DG**	**AP**	**ST**	**BS**	**AG**	**Users Accuracy (%)**
BR	48	5	1				2		85.71
BU		18					1		94.74
WA			37						100.00
DG		1		50		4		1	89.29
AP	1				50				98.04
ST	1					46			97.87
BS		1					35		97.22
AG							1	49	98.00
Producers accuracy (%)	96.00	72.00	97.37	100.00	100.00	92.00	89.74	98.00	
Overall accuracy (%) = 94.6									

Table S2. Confusion matrix for the SAR only classification. AG = agriculture, AP = alpine grassland, BR = bare, BU = built up, BS = bushes, DG = dry grassland, ST = steppe, WA = water.

Classified Data	**Reference Data**								
	BR	**BU**	**WA**	**DG**	**AP**	**ST**	**BS**	**AG**	**Users Accuracy (%)**
BR	49	4	2			1	2	1	83.05
BU		20	1						95.24
WA			29						100.00
DG			2	49	1	3		2	85.96
AP			1		48				97.96
ST	1		2	1		44		3	86.27
BS			1				37		97.37
AG		1			1	2		44	91.67
Producers accuracy (%)	98.00	80.00	76.32	98.00	96.00	88.00	94.87	88.00	
Overall accuracy (%) = 90.91									

Table S3. Confusion matrix for the OLI only classification. AG = agriculture, AP = alpine grassland, BR = bare, BU = built up, BS = bushes, DG = dry grassland, ST = steppe, WA = water.

Classified Data	**Reference Data**								
	BR	**BU**	**WA**	**DG**	**AP**	**ST**	**BS**	**AG**	**Users Accuracy (%)**
BR	42	4	1			3	1		82.35
BU	1	18					6		72.00
WA	5		37						88.10
DG	1			50		2		1	92.59
AP					49		1		98.00
ST		1				45			97.83
BS		1					24	1	92.31
AG	1	1			1		7	48	82.76
Producers accuracy (%)	84.00	72.00	97.37	100.00	98.00	90.00	61.54	96.00	
Overall accuracy (%) = 88.92									

Table S4. Cross-tabulated frequencies and McNemar's test p-values for the classification results (correct or incorrect) for the (**a**) SAROLI and OLI-only classifications; (**c**) SAROLI and SAR-only classifications, and; (**c**) SAR-only and OLI-only classifications.

a)	**SAROLI – OLI only**	OLI only	
	SAROLI	Incorrect	Correct
	Incorrect	9	10
	Correct	30	303
	p-value	0.003	
b)	**SAROLI – SAR only**	SAR only	
	SAROLI	Incorrect	Correct
	Incorrect	9	10
	Correct	23	310
	p-value	0.037	
c)	**OLI only – SAR only**	OLI only	
	SAR only	Incorrect	Correct
	Incorrect	1	31
	Correct	38	282
	p-value	0.470	

Table S5. Pearson product-moment correlation, p-values and statistical significance for *E. tancrei* and *M. gregalis* in relation to individual land cover class presence for the 50 m to 500 m nested buffers derived from the SAROLI land cover classification. ** denotes statistical significance at the 0.01 significance level, * denotes statistical significance at the 0.05 significance level.

		E. tancrei			*M. gregalis*		
Land Cover Class	**Buffer Radius (m)**	**Pearson's Correlation**	**p-Value**		**Pearson's Correlation**	**p-Value**	
Agriculture	50	0.409	<0.001	**	−0.206	0.045	*
Agriculture	100	0.438	<0.001	**	−0.187	0.070	
Agriculture	150	0.451	<0.001	**	−0.156	0.131	
Agriculture	200	0.414	<0.001	**	−0.158	0.126	
Agriculture	250	0.401	<0.001	**	−0.135	0.193	
Agriculture	300	0.391	<0.001	**	−0.131	0.206	
Agriculture	350	0.360	<0.001	**	−0.168	0.104	
Agriculture	400	0.370	<0.001	**	−0.181	0.080	
Agriculture	450	0.409	<0.001	**	−0.185	0.073	
Agriculture	500	0.412	<0.001	**	−0.165	0.111	
Alpine grassland	50	−0.093	0.371		−0.068	0.511	
Alpine grassland	100	−0.042	0.683		−0.051	0.627	
Alpine grassland	150	−0.034	0.742		−0.029	0.780	
Alpine grassland	200	−0.031	0.762		−0.030	0.775	
Alpine grassland	250	−0.060	0.563		−0.062	0.554	
Alpine grassland	300	−0.056	0.591		−0.053	0.610	
Alpine grassland	350	−0.084	0.421		−0.032	0.759	
Alpine grassland	400	−0.082	0.430		−0.046	0.656	
Alpine grassland	450	−0.090	0.387		−0.053	0.610	
Alpine grassland	500	−0.084	0.419		−0.051	0.621	

Bare	50	−0.109	0.293		0.022	0.832	
Bare	100	−0.194	0.059		−0.049	0.640	
Bare	150	−0.160	0.123		−0.070	0.502	
Bare	200	−0.144	0.164		−0.049	0.640	
Bare	250	−0.105	0.310		0.003	0.977	
Bare	300	−0.133	0.197		−0.030	0.775	
Bare	350	−0.141	0.174		−0.020	0.846	
Bare	400	−0.173	0.093		−0.013	0.900	
Bare	450	−0.116	0.264		0.057	0.585	
Bare	500	−0.151	0.145		0.063	0.542	
Built-up	50	−0.008	0.938		−0.089	0.390	
Built-up	100	−0.057	0.585		−0.101	0.329	
Built-up	150	−0.065	0.534		−0.103	0.319	
Built-up	200	−0.066	0.523		−0.095	0.361	
Built-up	250	−0.069	0.506		−0.092	0.375	
Built-up	300	−0.073	0.484		−0.096	0.357	
Built-up	350	−0.061	0.557		−0.114	0.271	
Built-up	400	−0.048	0.643		−0.112	0.280	
Built-up	450	−0.042	0.689		−0.088	0.396	
Built-up	500	−0.052	0.615		−0.115	0.266	
Bushes	50	−0.111	0.284		0.585	<0.001	**
Bushes	100	−0.072	0.488		0.544	<0.001	**
Bushes	150	−0.080	0.443		0.433	<0.001	**
Bushes	200	−0.064	0.538		0.386	<0.001	**
Bushes	250	−0.078	0.454		0.408	<0.001	**
Bushes	300	−0.103	0.321		0.462	<0.001	**
Bushes	350	−0.095	0.361		0.430	<0.001	**
Bushes	400	−0.082	0.432		0.391	<0.001	**
Bushes	450	−0.077	0.456		0.317	0.002	**
Bushes	500	−0.064	0.537		0.283	0.006	**
Dry grassland	50	−0.067	0.518		0.031	0.768	
Dry grassland	100	−0.094	0.365		0.015	0.884	
Dry grassland	150	−0.112	0.278		0.017	0.870	
Dry grassland	200	−0.073	0.480		0.025	0.811	
Dry grassland	250	−0.047	0.652		−0.019	0.856	
Dry grassland	300	−0.011	0.917		−0.041	0.693	
Dry grassland	350	0.051	0.624		−0.024	0.818	
Dry grassland	400	0.049	0.640		−0.011	0.914	
Dry grassland	450	0.018	0.862		0.023	0.822	
Dry grassland	500	0.034	0.745		0.058	0.574	
Steppe	50	−0.242	0.018	*	−0.032	0.758	
Steppe	100	−0.250	0.015	*	0.005	0.961	
Steppe	150	−0.238	0.020	*	0.015	0.882	

Steppe	200	−0.248	0.015	*	0.005	0.964	
Steppe	250	−0.247	0.016	*	0.025	0.812	
Steppe	300	−0.250	0.015	*	0.041	0.695	
Steppe	350	−0.267	0.009	**	0.045	0.663	
Steppe	400	−0.264	0.010	*	0.052	0.619	
Steppe	450	−0.263	0.010	*	0.026	0.801	
Steppe	500	−0.267	0.009	**	0.002	0.982	
Water	50	−0.066	0.528		0.022	0.834	
Water	100	−0.111	0.284		0.117	0.258	
Water	150	−0.074	0.476		0.341	0.001	**
Water	200	−0.070	0.501		0.422	<0.001	**
Water	250	−0.013	0.899		0.388	<0.001	**
Water	300	−0.019	0.858		0.318	0.002	**
Water	350	−0.037	0.724		0.319	0.002	**
Water	400	−0.050	0.631		0.287	0.005	**
Water	450	−0.063	0.546		0.271	0.008	**
Water	500	−0.072	0.487		0.246	0.016	*

Table S6. Pearson product-moment correlation, p-values and statistical significance for *E. tancrei* and *M. gregalis* in relation to individual land cover class presence for the 50 m to 500 m nested buffers derived from the SAR only land cover classification. ** denotes statistical significance at the 0.01 significance level, * denotes statistical significance at the 0.05 significance level.

		E. tancrei			*M. gregalis*		
Land Cover Class	**Buffer Radius (m)**	**Pearson's Correlation**	**p-Value**		**Pearson's Correlation**	**p-Value**	
Agriculture	50	0.410	<0.001	**	−0.197	0.055	
Agriculture	100	0.476	<0.001	**	−0.182	0.077	
Agriculture	150	0.479	<0.001	**	−0.189	0.066	
Agriculture	200	0.465	<0.001	**	−0.176	0.088	
Agriculture	250	0.471	<0.001	**	−0.167	0.106	
Agriculture	300	0.436	<0.001	**	−0.182	0.078	
Agriculture	350	0.424	<0.001	**	−0.195	0.058	
Agriculture	400	0.406	<0.001	**	−0.175	0.089	
Agriculture	450	0.396	<0.001	**	−0.169	0.101	
Agriculture	500	0.432	<0.001	**	−0.166	0.107	
Alpine grassland	50	−0.090	0.388		−0.069	0.505	
Alpine grassland	100	−0.028	0.789		−0.008	0.939	
Alpine grassland	150	−0.006	0.957		−0.014	0.895	
Alpine grassland	200	−0.003	0.977		−0.029	0.781	
Alpine grassland	250	−0.027	0.799		−0.020	0.848	
Alpine grassland	300	−0.035	0.733		−0.019	0.858	
Alpine grassland	350	−0.080	0.444		−0.022	0.829	
Alpine grassland	400	−0.096	0.355		−0.019	0.858	

Alpine grassland	450	−0.079	0.444		−0.009	0.932	
Alpine grassland	500	−0.068	0.515		−0.010	0.922	
Bare	50	−0.128	0.215		0.173	0.093	
Bare	100	−0.138	0.183		0.219	0.033	*
Bare	150	−0.141	0.172		0.136	0.187	
Bare	200	−0.142	0.171		0.148	0.152	
Bare	250	−0.091	0.382		0.177	0.086	
Bare	300	−0.129	0.214		0.112	0.280	
Bare	350	−0.168	0.103		0.085	0.415	
Bare	400	−0.190	0.065		0.102	0.326	
Bare	450	−0.155	0.134		0.187	0.070	
Bare	500	−0.134	0.194		0.169	0.101	
Built-up	50	−0.016	0.881		−0.063	0.547	
Built-up	100	−0.047	0.654		−0.067	0.519	
Built-up	150	−0.037	0.723		−0.079	0.447	
Built-up	200	−0.036	0.726		−0.112	0.281	
Built-up	250	−0.045	0.667		−0.093	0.372	
Built-up	300	−0.039	0.709		−0.097	0.349	
Built-up	350	−0.023	0.827		−0.116	0.265	
Built-up	400	−0.042	0.685		−0.098	0.343	
Built-up	450	−0.017	0.873		−0.080	0.444	
Built-up	500	−0.029	0.777		−0.114	0.270	
Bushes	50	−0.111	0.284		0.487	<0.001	**
Bushes	100	−0.099	0.341		0.331	0.001	**
Bushes	150	−0.070	0.497		0.317	0.002	**
Bushes	200	−0.048	0.645		0.279	0.006	**
Bushes	250	−0.081	0.433		0.321	0.002	**
Bushes	300	−0.064	0.535		0.368	<0.001	**
Bushes	350	−0.063	0.541		0.407	<0.001	**
Bushes	400	−0.070	0.501		0.401	<0.001	**
Bushes	450	−0.038	0.716		0.239	0.020	*
Bushes	500	−0.051	0.625		0.180	0.081	
Dry grassland	50	0.027	0.797		−0.025	0.810	
Dry grassland	100	−0.006	0.956		−0.076	0.466	
Dry grassland	150	−0.043	0.680		−0.079	0.444	
Dry grassland	200	−0.027	0.795		−0.034	0.742	
Dry grassland	250	−0.007	0.950		−0.071	0.492	
Dry grassland	300	0.020	0.847		−0.094	0.366	
Dry grassland	350	0.062	0.549		−0.075	0.469	
Dry grassland	400	0.082	0.431		−0.066	0.528	
Dry grassland	450	0.071	0.493		−0.067	0.517	
Dry grassland	500	0.055	0.595		−0.040	0.697	
Steppe	50	−0.266	0.009	**	−0.010	0.927	

Steppe	100	−0.259	0.011	*	0.043	0.681	
Steppe	150	−0.231	0.024	*	0.110	0.287	
Steppe	200	−0.254	0.013	*	0.049	0.637	
Steppe	250	−0.255	0.013	*	0.065	0.533	
Steppe	300	−0.233	0.023	*	0.101	0.332	
Steppe	350	−0.233	0.023	*	0.088	0.395	
Steppe	400	−0.222	0.030	*	0.064	0.536	
Steppe	450	−0.231	0.024	*	0.067	0.516	
Steppe	500	−0.232	0.024	*	0.067	0.519	
Water	50	−0.082	0.431		0.043	0.679	
Water	100	−0.113	0.274		0.153	0.140	
Water	150	−0.166	0.108		0.208	0.043	*
Water	200	−0.112	0.280		0.442	<0.001	**
Water	250	−0.088	0.399		0.257	0.012	*
Water	300	−0.105	0.309		0.303	0.003	**
Water	350	−0.113	0.274		0.393	<0.001	**
Water	400	−0.125	0.227		0.349	0.001	**
Water	450	−0.111	0.285		0.238	0.020	*
Water	500	−0.101	0.332		0.207	0.044	*

Table S7. Pearson product-moment correlation, p-values and statistical significance for *E. tancrei* and *M. gregalis* in relation to individual land cover class presence for the 50 m to 500 m nested buffers derived from the OLI only land cover classification. ** denotes statistical significance at the 0.01 significance level, * denotes statistical significance at the 0.05 significance level.

		E. tancrei			*M. gregalis*		
Land Cover Class	**Buffer Radius (m)**	**Pearson's Correlation**	**p-Value**		**Pearson's Correlation**	**p-Value**	
Agriculture	50	0.491	<0.001	**	−0.149	0.149	
Agriculture	100	0.512	<0.001	**	−0.225	0.029	*
Agriculture	150	0.529	<0.001	**	−0.214	0.038	*
Agriculture	200	0.491	<0.001	**	−0.206	0.046	*
Agriculture	250	0.469	<0.001	**	−0.161	0.119	
Agriculture	300	0.468	<0.001	**	−0.191	0.064	
Agriculture	350	0.458	<0.001	**	−0.203	0.048	*
Agriculture	400	0.430	<0.001	**	−0.210	0.041	*
Agriculture	450	0.465	<0.001	**	−0.197	0.056	
Agriculture	500	0.450	<0.001	**	−0.206	0.045	*
Alpine grassland	50	−0.055	0.594		−0.044	0.675	
Alpine grassland	100	−0.032	0.759		−0.065	0.533	
Alpine grassland	150	−0.038	0.716		−0.031	0.767	
Alpine grassland	200	−0.024	0.820		0.026	0.803	
Alpine grassland	250	−0.038	0.718		−0.054	0.606	
Alpine grassland	300	−0.020	0.849		−0.023	0.826	
Alpine grassland	350	−0.063	0.546		−0.021	0.841	
Alpine grassland	400	−0.040	0.702		0.017	0.868	
Alpine grassland	450	−0.071	0.495		−0.002	0.986	
Alpine grassland	500	−0.106	0.306		0.007	0.946	
Bare	50	−0.158	0.126		−0.061	0.556	
Bare	100	−0.144	0.163		0.043	0.677	
Bare	150	−0.156	0.132		−0.019	0.857	
Bare	200	−0.161	0.120		0.007	0.945	
Bare	250	−0.201	0.051		−0.036	0.731	
Bare	300	−0.243	0.018	*	−0.022	0.835	
Bare	350	−0.220	0.032	*	−0.028	0.788	
Bare	400	−0.198	0.055		−0.006	0.952	
Bare	450	−0.181	0.079		0.051	0.624	
Bare	500	−0.260	0.011	*	0.108	0.296	
Built-up	50	−0.012	0.906		−0.048	0.643	
Built-up	100	−0.078	0.454		−0.076	0.463	
Built-up	150	−0.097	0.351		−0.153	0.138	
Built-up	200	−0.042	0.684		−0.080	0.443	
Built-up	250	−0.029	0.781		−0.026	0.805	
Built-up	300	−0.030	0.776		−0.026	0.801	
Built-up	350	−0.041	0.690		−0.031	0.765	

Built-up	400	−0.036	0.729		−0.026	0.804	
Built-up	450	−0.013	0.897		−0.003	0.976	
Built-up	500	−0.005	0.965		−0.030	0.776	
Bushes	50	−0.043	0.682		0.420	<0.001	**
Bushes	100	−0.011	0.918		0.407	<0.001	**
Bushes	150	−0.027	0.795		0.488	<0.001	**
Bushes	200	−0.052	0.616		0.454	<0.001	**
Bushes	250	−0.061	0.559		0.357	<0.001	**
Bushes	300	−0.082	0.430		0.504	<0.001	**
Bushes	350	−0.055	0.599		0.432	<0.001	**
Bushes	400	−0.041	0.696		0.250	0.015	*
Bushes	450	−0.031	0.763		0.171	0.097	
Bushes	500	−0.008	0.938		0.203	0.048	*
Dry grassland	50	−0.173	0.093		0.042	0.685	
Dry grassland	100	−0.181	0.079		0.084	0.417	
Dry grassland	150	−0.193	0.060		0.073	0.484	
Dry grassland	200	−0.140	0.175		0.033	0.751	
Dry grassland	250	−0.092	0.378		−0.009	0.932	
Dry grassland	300	−0.083	0.423		−0.033	0.754	
Dry grassland	350	−0.025	0.809		−0.016	0.876	
Dry grassland	400	−0.020	0.847		−0.014	0.890	
Dry grassland	450	−0.039	0.709		0.024	0.819	
Dry grassland	500	0.011	0.916		0.044	0.675	
Steppe	50	−0.261	0.010	*	0.007	0.947	
Steppe	100	−0.261	0.011	*	0.057	0.584	
Steppe	150	−0.260	0.011	*	0.067	0.519	
Steppe	200	−0.272	0.008	**	0.056	0.592	
Steppe	250	−0.276	0.007	**	0.073	0.479	
Steppe	300	−0.275	0.007	**	0.089	0.390	
Steppe	350	−0.285	0.005	**	0.073	0.483	
Steppe	400	−0.282	0.006	**	0.090	0.385	
Steppe	450	−0.282	0.006	**	0.057	0.584	
Steppe	500	−0.281	0.006	**	0.042	0.687	
Water	50	−0.030	0.773		−0.043	0.680	
Water	100	−0.076	0.466		−0.077	0.457	
Water	150	−0.031	0.763		0.217	0.035	*
Water	200	−0.070	0.501		0.328	0.001	**
Water	250	−0.001	0.994		0.328	0.001	**
Water	300	−0.002	0.983		0.244	0.017	*
Water	350	−0.046	0.658		0.335	0.001	**
Water	400	−0.064	0.541		0.295	0.004	**
Water	450	−0.069	0.506		0.250	0.015	*
Water	500	−0.072	0.490		0.227	0.027	*

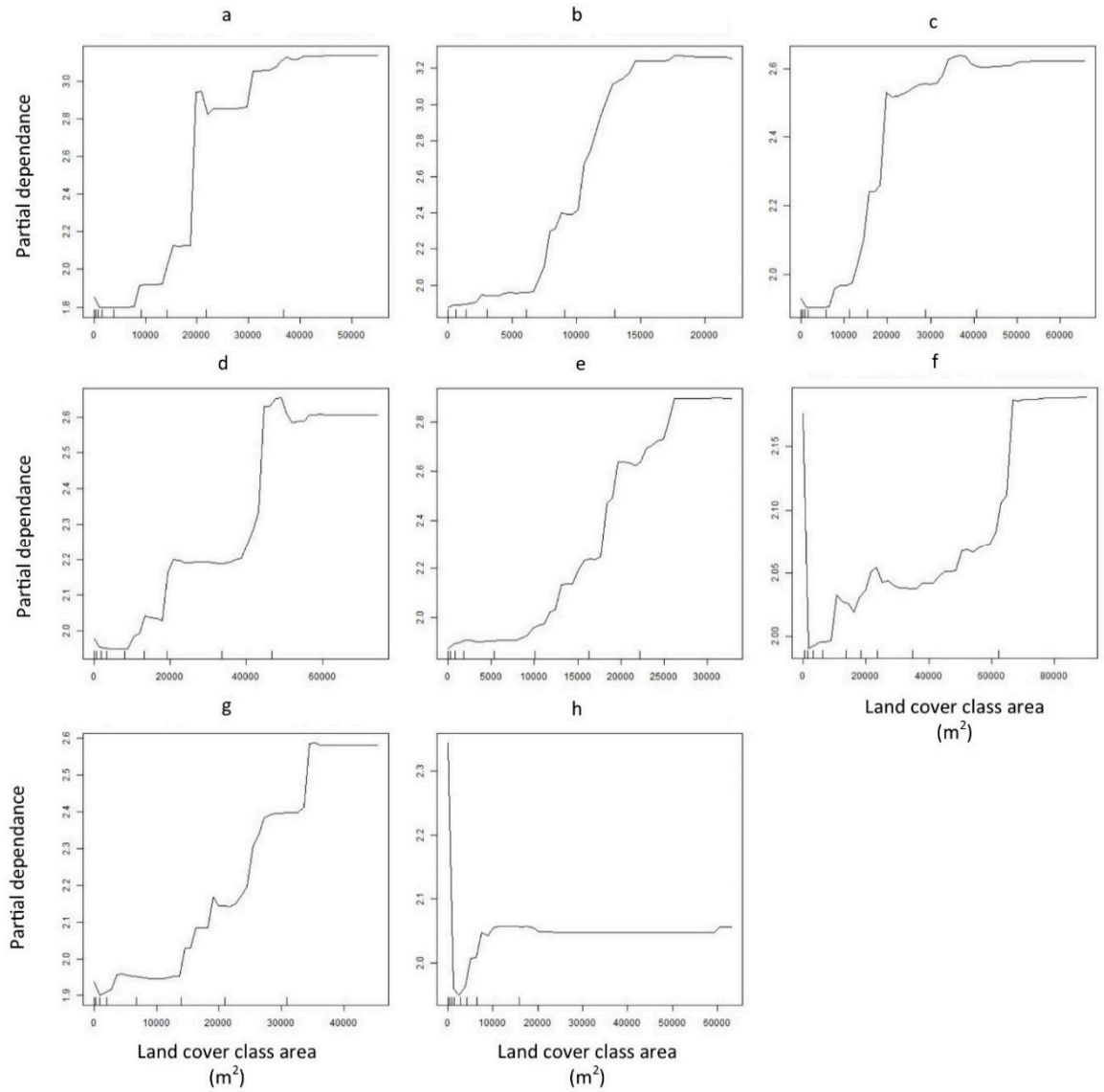

Figure S1. Random forest partial dependence plots for the statistically significant (as shown by random forest permutation analysis) nested land cover variables derived from the SAR only land cover classification, and *E. tancrei* relative density. Presented in order of variable importance, (**a**) = agriculture 250 m, (**b**) = agriculture 100 m, (**c**) = agriculture 300 m, (**d**) = agriculture 350 m, (**e**) = agriculture 150 m, (**f**) = agriculture 450 m, (**g**) = agriculture 200 m, (**h**) = built-up 500 m.

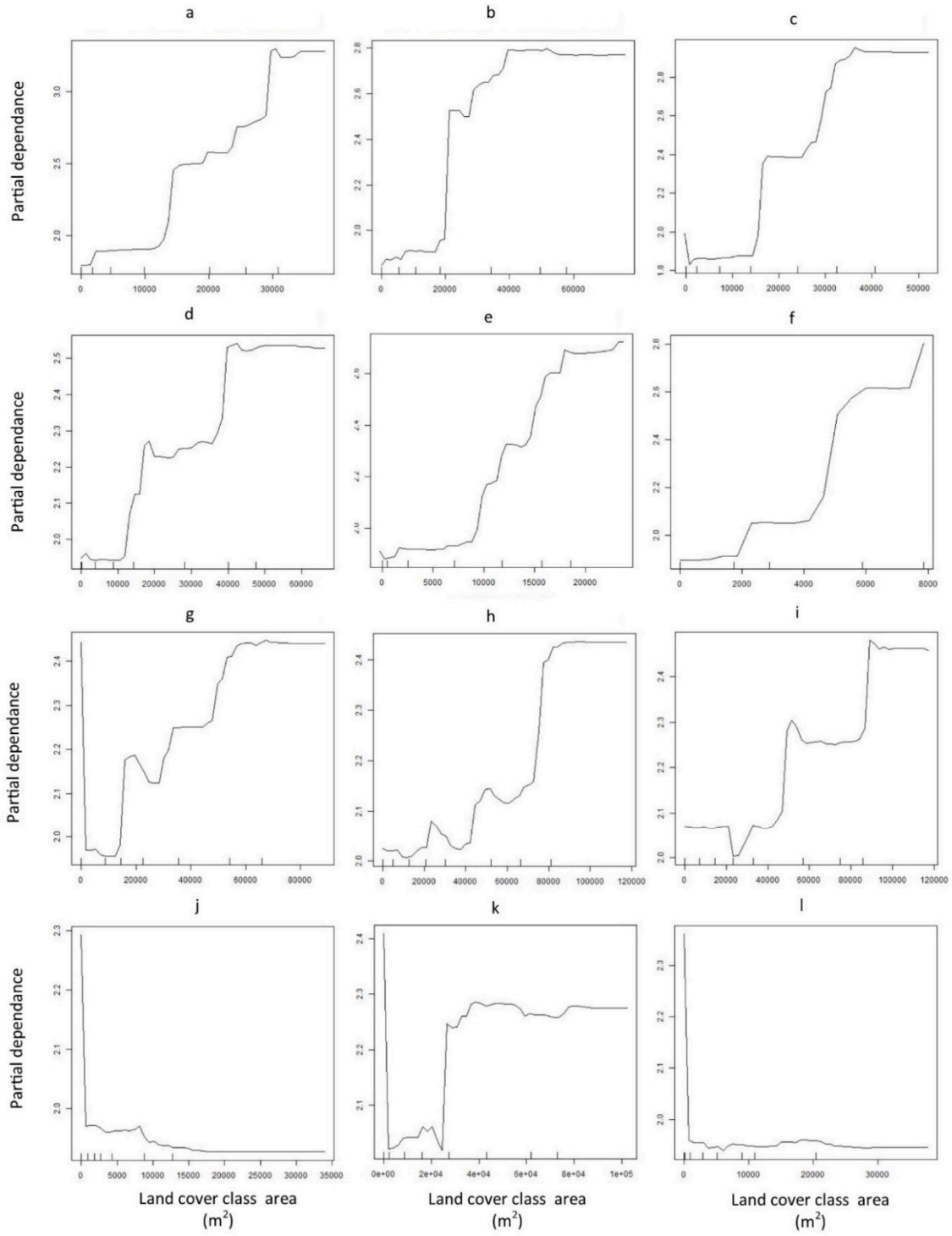

Figure S2. Random forest partial dependence plots for the statistically significant (as shown by random forest permutation analysis) nested land cover variables derived from the OLI only land cover classification, and *E. tancrei* relative density. Presented in order of variable importance, (**a**) = agriculture 150 m, (**b**) = agriculture 300 m, (**c**) = agriculture 200 m, (**d**) = agriculture 250 m, (**e**) = agriculture 100 m, (**f**) = agriculture 50 m, (**g**) = agriculture 350 m, (**h**) = agriculture 450 m, (**i**) = agriculture 500 m, (**j**) = built-up 400 m, (**k**) = agriculture 400 m, (**l**) = built-up 500 m.