Long-Term Land Use/Land Cover Change Assessment of the Kilombero Catchment in Tanzania Using Random Forest Classification and Robust Change Vector Analysis

Frank Thonfeld [1,2,*], Stefanie Steinbach [3], Javier Muro [4] and Fridah Kirimi [5]

[1] German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Münchener Straße 20, 82234 Weßling, Germany; frank.thonfeld@dlr.de (F.T.)
[2] Department of Remote Sensing, University of Würzburg, Oswald-Külpe-Str. 86, 97074 Würzburg, Germany
[3] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, 7514 AE Enschede, The Netherlands; s.steinbach@utwente.nl (S.S.)
[4] Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn, 53113 Bonn, Germany; jmuro@uni-bonn.de (J.M.)
[5] Department of Geomatic Engineering and Geospatial Information Systems, Jomo Kenyatta University of Agriculture and Technology, P.O. Box 62000 Nairobi, Kenya; fkirimi@jkuat.ac.ke (F.K.)
* Correspondence: frank.thonfeld@dlr.de; Tel.: +49-8153-28-4306 (F.T.)

Supplementary

Table S1. Random forest model performance for the 1974 classification (number of trees: 500, no. of variables tried at each split: 6, out-of-bag (OOB) estimate of error rate: 3.28%).

	built-up	closed woodland	upland agriculture	grassland	montane forest	open woodland	rice	savanna	swamp	water	class.error
built-up	10	0	2	1	0	4	2	3	0	0	0.55
closed woodland	0	260	0	0	1	4	0	0	0	0	0.02
upland agriculture	0	0	72	2	0	1	0	2	0	0	0.06
grassland	0	0	1	177	0	0	0	2	0	0	0.02
montane forest	0	0	0	0	116	0	0	0	0	0	0.00
open woodland	0	5	1	1	1	372	0	2	1	0	0.03
rice	0	0	0	1	0	0	23	1	0	0	0.08
savanna	0	0	1	3	0	3	0	138	1	0	0.05
swamp	0	1	0	1	0	1	0	1	67	0	0.06
water	0	1	0	14	0	1	0	0	1	42	0.29

Table S2. Random forest model performance for the 1994 classification (number of trees: 500, no. of variables tried at each split: 24, OOB estimate of error rate: 6.67%).

	built-up	closed woodland	cropland	grassland	montane forest	open woodland	rice	savanna	swamp	water	class.error
built-up	17	0	9	2	0	0	1	2	1	0	0.47
closed woodland	0	241	0	0	1	11	0	1	0	0	0.05
upland agriculture	0	0	85	2	0	1	0	2	0	0	0.06
grassland	0	0	3	168	0	1	0	2	0	1	0.04
montane forest	0	0	0	0	119	1	0	0	0	0	0.01
open woodland	0	13	2	2	2	326	0	4	0	0	0.07
rice	0	0	4	2	0	0	33	2	0	0	0.20
savanna	0	2	2	5	0	8	1	157	1	0	0.11
swamp	0	0	0	1	0	1	0	2	41	0	0.09
water	0	0	1	12	0	2	0	0	1	46	0.26

Table S3. Random forest model performance for the 2004 classification (number of trees: 500, no. of variables tried at each split: 32, OOB estimate of error rate: 6.21%).

	built-up	closed woodland	upland agriculture	grassland	montane forest	open woodland	rice	savanna	swamp	teak	water	class.error
built-up	20	0	7	1	0	4	2	4	0	0	0	0.47
closed woodland	0	186	0	0	2	9	0	0	0	0	0	0.06
upland agriculture	0	0	100	2	0	6	1	2	0	0	0	0.10
grassland	0	0	0	171	0	0	1	1	1	0	0	0.02
montane forest	0	0	0	0	118	1	0	0	0	0	0	0.01
open woodland	0	8	3	2	1	311	1	3	1	0	0	0.06
rice	0	0	3	3	0	0	62	1	0	0	0	0.10
savanna	0	0	2	4	0	9	2	130	1	0	0	0.12
swamp	0	0	0	2	0	1	0	0	42	0	0	0.07
teak	0	0	0	0	0	0	0	0	0	50	0	0.00
water	0	0	1	3	0	1	1	0	0	0	57	0.10

Table S4. Random forest model performance for the 2014 classification (number of trees: 500, no. of variables tried at each split: 32, OOB estimate of error rate: 6.38%).

	built-up	closed woodland	upland agriculture	grassland	montane forest	open woodland	rice	savanna	swamp	teak	water	class.error
built-up	24	0	9	1	0	2	3	1	0	0	0	0.40
closed woodland	0	189	1	0	1	8	0	0	0	0	0	0.05
upland agriculture	0	0	155	1	0	8	7	0	0	0	0	0.09
grassland	0	0	1	118	0	0	5	0	0	0	0	0.05
montane forest	0	0	0	0	118	1	0	0	0	0	0	0.01
open woodland	0	7	4	1	1	269	3	1	0	0	0	0.06
rice	0	0	5	3	0	1	157	0	0	0	0	0.05
savanna	0	0	4	0	0	5	3	67	0	0	0	0.15
swamp	0	0	0	1	0	2	1	0	45	0	0	0.08
teak	0	0	0	0	0	1	0	0	0	49	0	0.02
water	0	1	0	7	0	1	1	0	0	0	51	0.16

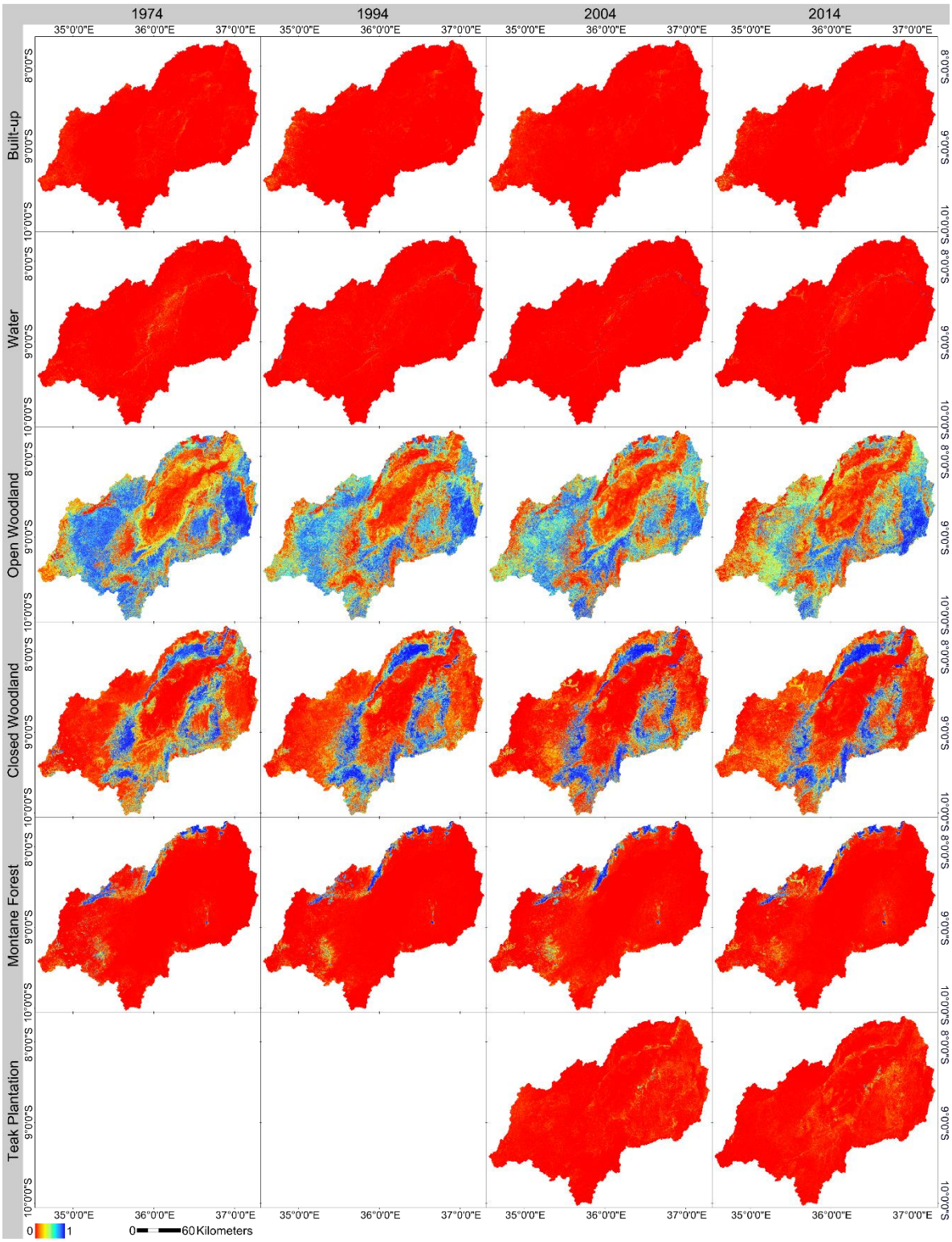

Figure S1. Classification probabilities per pixel and per class. Higher values indicate high probability of belonging to a certain class. Probabilities do not translate in assignment to a certain class if another class has higher probabilities. Therefore, it may happen that two classes have probabilities close to 0.5 and it is worth looking in the probability maps.

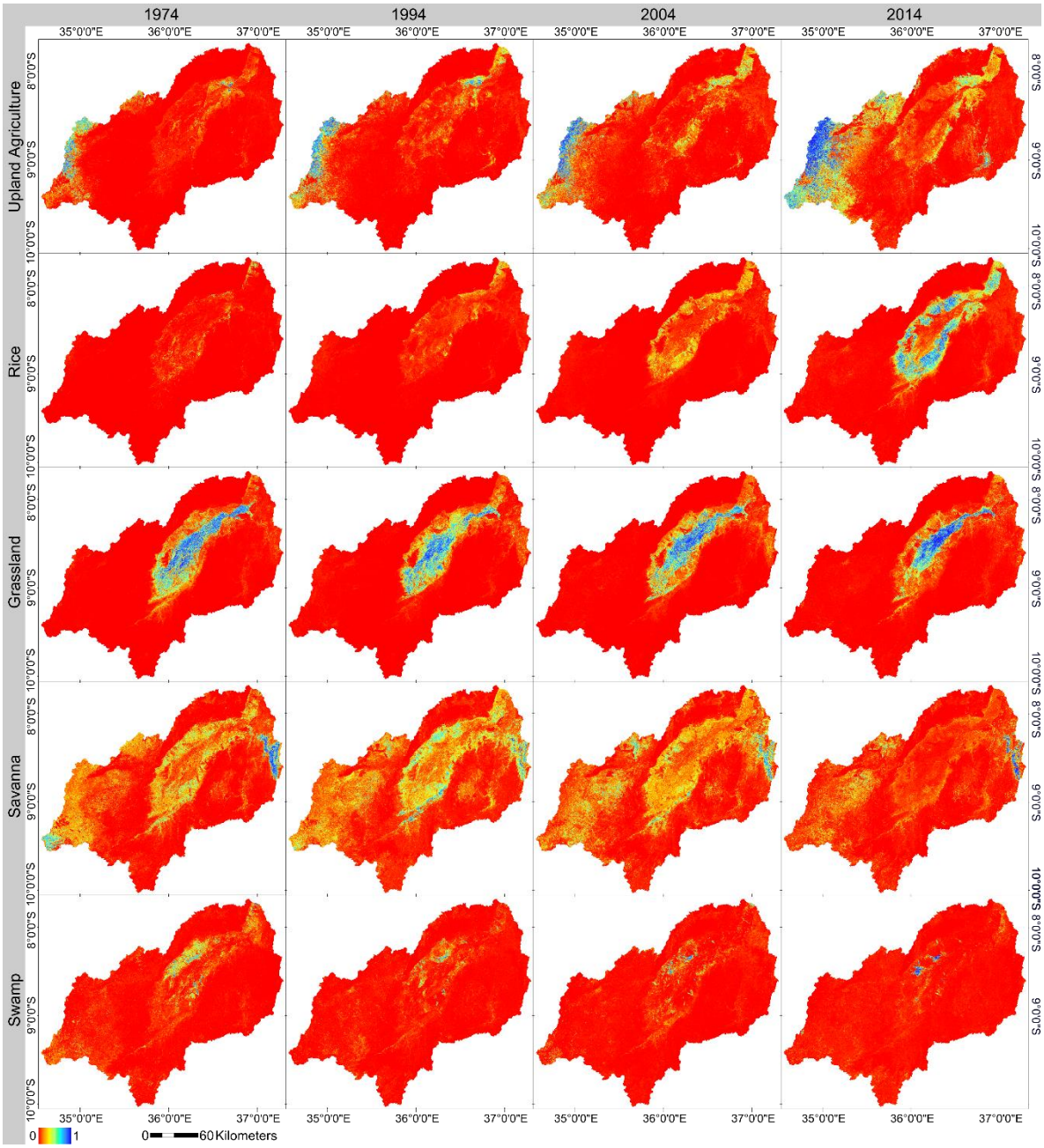

Figure S1 (cont.)

Table S5. From-to-LULC changes between 1974 and 1994 (in ha).

	1994 Built-up	1994 Closed Woodland	1994 Upland Agriculture	1994 Grassland	1994 Montane Forest	1994 Open Woodland	1994 Rice	1994 Savanna	1994 Swamp	1994 Teak	1994 Water
1974 Built-up	45	51	357	15	3	105	3	64	4	0	3
1974 Closed Woodland	57	636,601	4,999	1,581	3,164	240,999	2,341	11,551	1,693	0	418
1974 Upland Agriculture	549	949	71,645	7,552	571	17,164	466	12,582	1,296	0	134
1974 Grassland	65	2,721	13,810	199,992	0	14,564	2,761	61,796	8,962	0	2,326
1974 Montane Forest	10	8,577	1,195	0	88,083	61,932	0	713	974	0	35
1974 Open Woodland	611	331,589	58,659	11,746	41,401	1,554,343	3,890	120,173	7,885	0	1,370
1974 Rice	36	238	3,491	3,965	0	811	3,236	5,135	581	0	12
1974 Savanna	340	4,635	27,740	46,421	1,208	75,055	3,262	115,670	6,608	0	672
1974 Swamp	30	4,011	7,687	23,073	203	16,902	3,778	32,326	12,620	0	262
1974 Teak	0	0	0	0	0	0	0	0	0	0	0
1974 Water	10	242	101	1,478	32	1,088	8	422	98	0	3,283

Table S6. From-to-LULC changes between 1994 and 2004 (in ha).

	2004 Built-up	2004 Closed Woodland	2004 Upland Agriculture	2004 Grassland	2004 Montane Forest	2004 Open Woodland	2004 Rice	2004 Savanna	2004 Swamp	2004 Teak	2004 Water
1994 Built-up	384	7	952	18	1	204	18	152	10	3	4
1994 Closed Woodland	9	696,752	1,868	1,829	7,124	272,885	1,011	3,022	1,680	2,614	824
1994 Upland Agriculture	936	696	118,759	13,943	211	32,979	7,303	12,345	2,238	180	94
1994 Grassland	61	379	19,443	203,220	0	6,284	25,050	29,319	8,874	27	3,165
1994 Montane Forest	1	8,767	209	0	101,982	23,237	0	222	249	0	27
1994 Open Woodland	127	143,652	52,445	10,063	38,259	1,617,494	4,767	106,364	7,433	2,072	362
1994 Rice	30	26	5,546	2,592	0	1,485	3,636	5,010	1,301	108	11
1994 Savanna	125	1,664	37,368	71,240	163	95,873	28,783	109,267	15,120	442	387
1994 Swamp	7	764	3,263	11,283	664	6,313	2,318	4,095	11,521	38	455
1994 Teak	0	0	0	0	0	0	0	0	0	0	0
1994 Water	2	116	71	875	20	384	33	329	287	0	6,398

Table S7. From-to-LULC changes between 2004 and 2014 (in ha).

	2014 Built-up	2014 Closed Woodland	2014 Upland Agriculture	2014 Grassland	2014 Montane Forest	2014 Open Woodland	2014 Rice	2014 Savanna	2014 Swamp	2014 Teak	2014 Water
2004 Built-up	651	2	763	32	1	85	133	11	2	1	1
2004 Closed Woodland	32	658,585	5,066	304	5,188	178,773	2,213	381	984	1,234	65
2004 Upland Agriculture	2,223	2,461	163,860	4,741	313	14,622	47,116	3,943	296	292	56
2004 Grassland	8	270	13,153	168,799	0	14,184	109,917	2,450	4,763	20	1,498
2004 Montane Forest	32	9,239	4,963	0	100,351	33,544	0	234	29	0	33
2004 Open Woodland	501	208,187	283,443	6,794	20,973	1,443,883	35,406	51,605	3,705	2,199	443
2004 Rice	26	88	4,949	8,136	0	4,340	54,560	273	412	56	80
2004 Savanna	207	2,001	46,380	19,830	398	85,190	67,576	45,483	2,566	80	412
2004 Swamp	20	1,869	3,552	11,240	263	7,347	11,741	725	11,719	31	206
2004 Teak	0	552	806	28	0	1,740	401	88	6	1,863	0
2004 Water	0	288	61	2,134	6	1,009	198	364	1,778	2	5,886

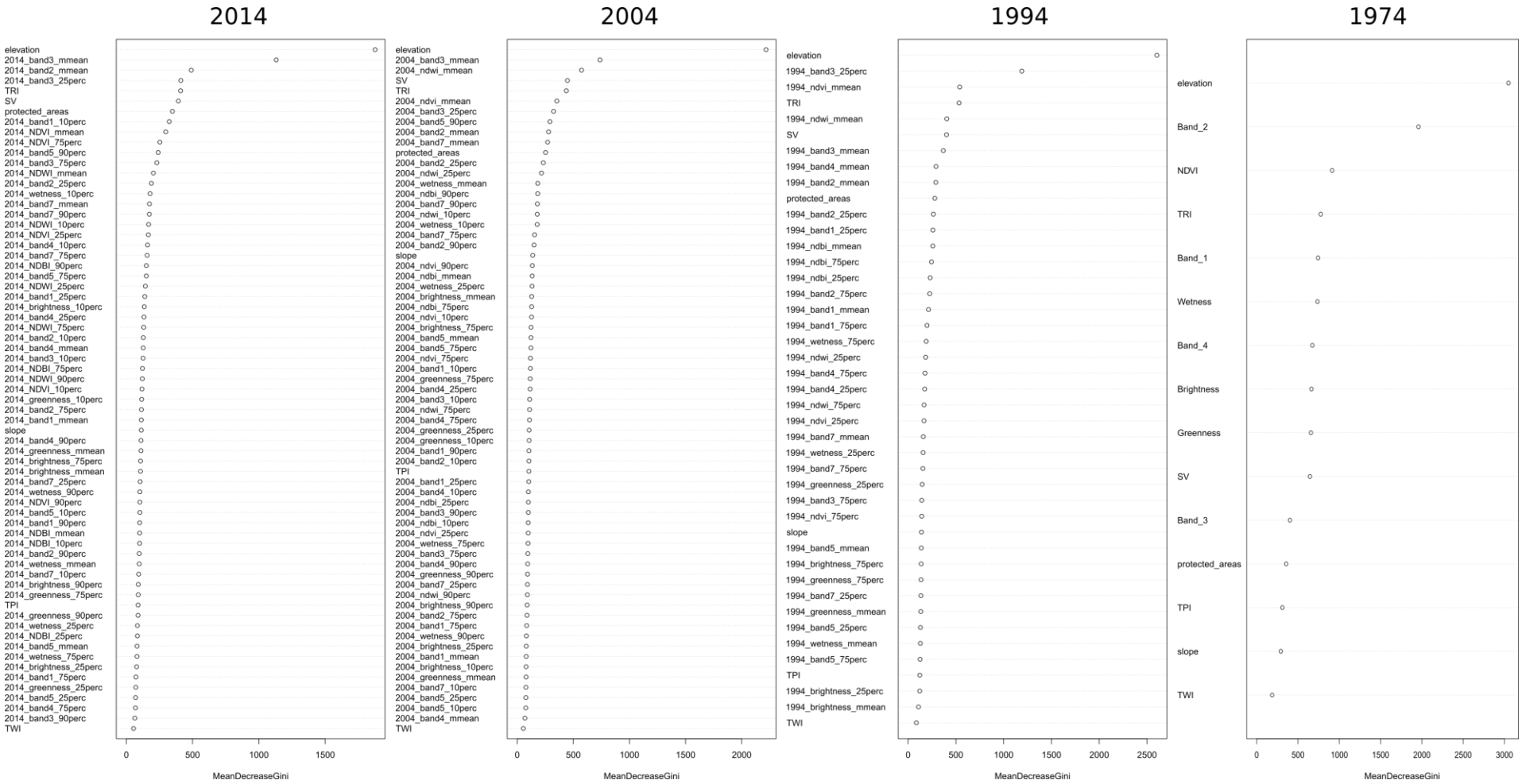

Figure S2. Variable importance for the classifications of 1974, 1994, 2004, and 2014.