Estimating Forest Gross Primary Production Using Machine Learning, Light Use Efficiency Model, and Global Eddy Covariance Data

Zhenkun Tian[1], Yingying Fu[2*], Tao Zhou[3,4], Chuixiang Yi[5,6], Eric Kutter[7], Qin Zhang[8], Nir Y Krakauer[9]

[1]School of Computer Science, China University of Labor Relations, Beijing 100048, China
[2]School of Mathematics and Statistics, Beijing Technology and Business University, Beijing 100048, China
[3]State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
[4]Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[5]School of Earth and Environmental Sciences, Queens College, City University of New York, Flushing, NY 11367, USA
[6]Earth and Environmental Sciences Department, Graduate Center, City University of New York, New York, NY 10016, USA
[7]Barry Commoner Center for Health and the Environment, Queens College, City University of New York, NY 11367, USA
[8]Institution of Water and Environment Research, Dalian University of Technology, Dalian 116024, China
[9]Department of Civil Engineering and NOAA-CREST, The City College of New York, New York, NY, USA

Contents of this file

Figure S1. The scatter plots of the square of the correlation coefficients (R^2), root mean square error (RMSE, gC m^{-2}day^{-1}), and relative predictive error (RPE) across evergreen broadleaf forest (EBF) between daily site-derived GPP (gC m^{-2}day^{-1}) at flux sites and the estimates from five individual models, MODIS GPP product, and machine learning methods (SVM and RF).
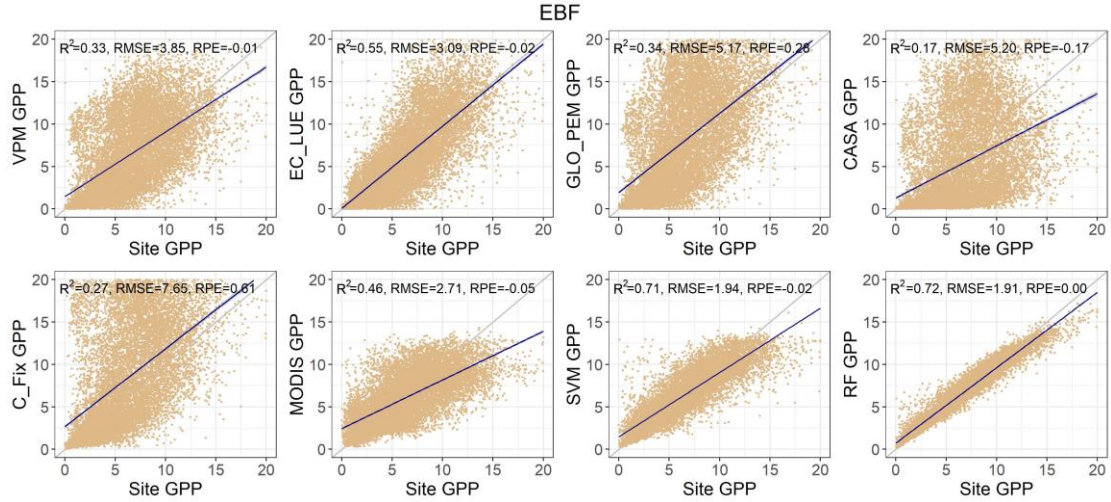

Figure S2. The scatter plots of the square of the correlation coefficients (R^2), root mean square error (RMSE, gC m^{-2}day^{-1}), and relative predictive error (RPE) across evergreen needleleaf forest (ENF) between daily site-derived GPP (gC m^{-2}day^{-1}) at flux sites and the estimates from five individual models, MODIS GPP product, and machine learning methods (SVM and RF).

MF

Figure S3. The scatter plots of the square of the correlation coefficients (R^2), root mean square error (RMSE, gC m^{-2}day^{-1}), and relative predictive error (RPE) across mixed forest (MF) between daily site-derived GPP (gC m^{-2}day^{-1}) at flux sites and the estimates from five individual models, MODIS GPP product, and machine learning methods (SVM and RF).
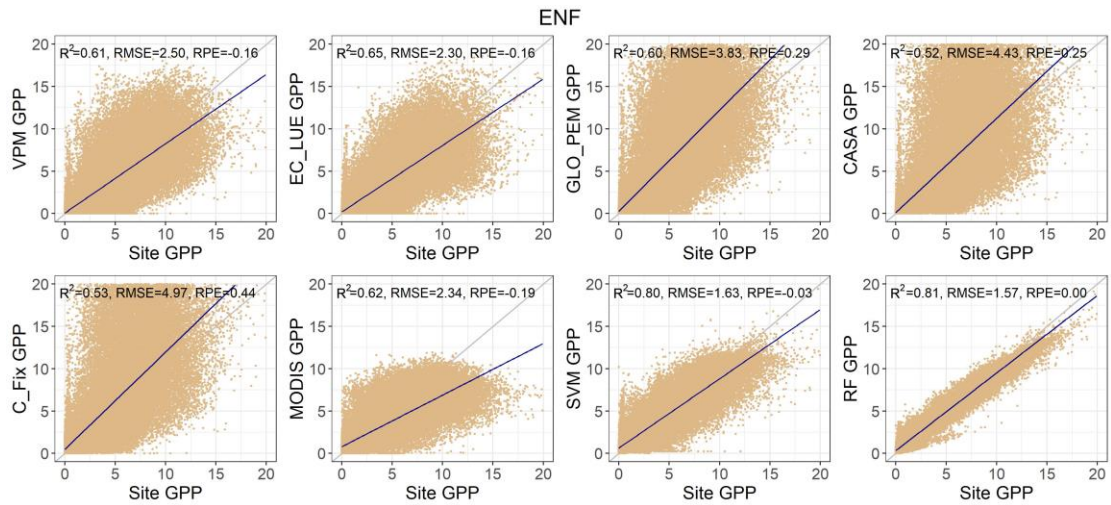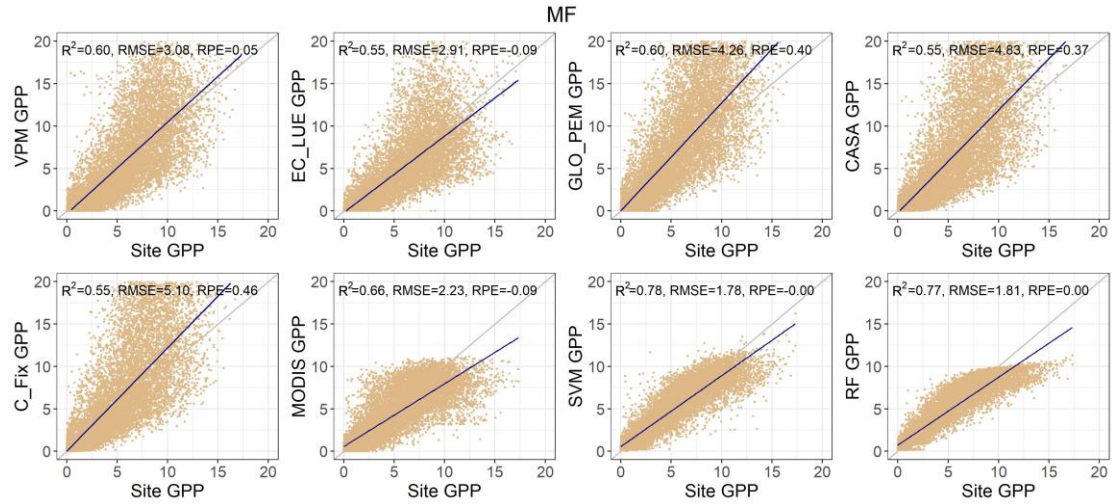
Table S1. The information of selected 45 forest FLUXNET2015 sites.

No	SITE_ID	Latitude	Longitude	Forest type	Observation years	Climate zone
1	AU-Tum	-35.6566	148.1517	EBF	14	Cfb
2	AU-Wom	-37.4222	144.0944	EBF	5	Cfb
3	BE-Bra	51.3076	4.5198	MF	19	Cfb
4	BE-Vie	50.3049	5.9981	MF	19	Cfb
5	BR-Sa3	-3.018	-54.9714	EBF	5	Am
6	CA-Gro	48.2167	-82.1556	MF	12	Dfb
7	CA-Obs	53.9872	-105.1178	ENF	14	Dfc
8	CA-Qfo	49.6925	-74.3421	ENF	8	Dfc
9	CA-SF2	54.2539	-105.8775	ENF	5	Dfc
10	CA-TP1	42.6609	-80.5595	ENF	13	Dfb
11	CA-TP2	42.7744	-80.4588	ENF	6	Dfb
12	CA-TP3	42.7068	-80.3483	ENF	13	Dfb
13	CA-TP4	42.7102	-80.3574	ENF	13	Dfb
14	CH-Dav	46.8153	9.8559	ENF	18	ET
15	DE-Hai	51.0792	10.4522	DBF	13	Cfb
16	DE-Lkb	49.0996	13.3047	ENF	5	Cfb
17	DE-Lnf	51.3282	10.3678	DBF	11	Cfb
18	DE-Obe	50.7867	13.7213	ENF	7	Cfb
19	DE-Tha	50.9626	13.5651	ENF	19	Cfb
20	DK-Sor	55.4859	11.6446	DBF	19	Cfb
21	FI-Hyy	61.8474	24.2948	ENF	19	Dfc
22	FI-Sod	67.3624	26.6386	ENF	14	Dfc
23	FR-LBr	44.7171	-0.7693	ENF	13	Cfb
24	FR-Pue	43.7413	3.5957	EBF	15	Csa
25	IT-Col	41.8494	13.5881	DBF	19	Cfa
26	IT-Cpz	41.7052	12.3761	EBF	13	Csa
27	IT-Lav	45.9562	11.2813	ENF	12	Cfb
28	IT-Ren	46.5869	11.4337	ENF	16	Dfc

No	SITE_ID	Latitude	Longitude	Forest type	Observation years	Climate zone
29	IT-Ro2	42.3903	11.9209	DBF	11	Csa
30	IT-SRo	43.7279	10.2844	ENF	14	Csa
31	NL-Loo	52.1666	5.7436	ENF	19	Cfb
32	RU-Fyo	56.4615	32.9221	ENF	17	Dfb
33	US-Blo	38.8953	-120.6328	ENF	11	Csa
34	US-GBT	41.3658	-106.2397	ENF	8	Dfc
35	US-GLE	41.3665	-106.2399	ENF	11	Dfc
36	US-Me2	44.4523	-121.5574	ENF	13	Csb
37	US-Me4	44.4992	-121.6224	ENF	5	Csb
38	US-Me6	44.3233	-121.6078	ENF	5	Csb
39	US-MMS	39.3232	-86.4131	DBF	16	Cfa
40	US-NR1	40.0329	-105.5464	ENF	17	Dfc
41	US-Oho	41.5545	-83.8438	DBF	10	Dfa
42	US-Prr	65.1237	-147.4876	ENF	5	Dfc
43	US-Syv	46.242	-89.3477	MF	14	Dfb
44	US-WCr	45.8059	-90.0799	DBF	16	Dfb
45	ZM-Mon	-15.4391	23.2525	DBF	10	Aw

Table S2. The Köppen-Geiger climate classes.

1st	2nd	3rd	Description
A			Tropical
	f		- Rainforest
	m		- Monsoon
	w		- Savannah
B			Arid
	w		- Desert
	s		- Steppe
		h	- Hot
		k	- Cold
C			Temperate
	s		- Dry summer
	w		- Dry winter
	f		- Without dry season
		a	- Hot summer
		b	- Warm summer
		c	- Cold summer
D			Cold
	s		- Dry summer
	w		- Dry winter
	f		- Without dry season
		a	- Hot summer
		b	- Warm summer
		c	- Cold summer
		c	- Very cold winter
			Polar
	T		- Tundra
	F		- Frost

Table S3. The summary of the statistics (R^2, RMSE (gC m^{-2} day^{-1}), and RPE) between modeled and FLUXNET2015 GPP across the 45 forest sites.

Sites	CASA			C_Fix			GLO_PEM			VPM			EC_LUE			MODIS			SVM			RF		
	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE
AU-Tum	0.31	6.09	-0.53	0.47	5.82	0.10	0.49	4.57	-0.08	0.50	4.20	-0.30	0.54	3.82	-0.15	0.35	3.56	-0.17	0.62	2.51	-0.07	0.93	1.18	-0.02
AU-Wom	0.28	4.54	-0.32	0.52	6.86	0.54	0.57	4.79	0.28	0.59	3.23	0.03	0.57	3.58	0.16	0.52	2.21	0.11	0.77	1.64	0.17	0.94	0.87	0.10
BE-Bra	0.67	4.01	0.35	0.68	4.17	0.41	0.73	3.44	0.38	0.72	2.26	0.00	0.76	1.86	-0.22	0.70	1.83	-0.05	0.83	1.38	0.08	0.88	1.24	0.11
BE-Vie	0.64	3.44	-0.05	0.65	3.61	0.03	0.71	3.00	0.02	0.72	2.56	-0.24	0.73	2.62	-0.35	0.74	2.41	-0.28	0.88	1.44	-0.10	0.90	1.37	-0.10
BR-Sa3	0.01	4.96	0.37	0.07	6.13	0.60	0.02	4.76	0.39	0.02	2.89	0.09	0.04	3.59	0.26	0.01	2.75	-0.10	0.33	1.35	-0.01	0.90	0.65	0.00
CA-Gro	0.72	4.75	0.65	0.73	5.08	0.80	0.79	3.97	0.63	0.80	2.58	0.27	0.84	1.92	0.13	0.86	1.34	-0.01	0.90	1.13	-0.04	0.91	1.15	-0.07
CA-Obs	0.42	0.39	4.03	0.39	0.75	9.88	0.42	0.56	4.30	0.40	0.42	3.06	0.52	0.34	2.61	0.73	0.23	3.27	0.50	0.58	7.58	0.60	0.31	3.72
CA-Qfo	0.67	2.99	0.57	0.67	3.48	0.83	0.74	2.95	0.67	0.74	1.35	0.03	0.83	0.91	-0.04	0.80	1.18	0.28	0.90	0.76	0.17	0.98	0.37	0.10
CA-SF2	0.71	2.69	-0.11	0.70	2.91	-0.01	0.76	2.53	-0.03	0.77	2.21	-0.43	0.82	1.83	-0.28	0.79	1.87	-0.29	0.87	1.68	-0.28	0.97	0.82	-0.14
CA-TP1	0.45	6.25	1.18	0.45	7.50	1.56	0.49	5.27	1.06	0.47	2.86	0.25	0.59	2.05	0.13	0.51	2.39	0.29	0.63	1.69	0.11	0.92	0.80	0.10
CA-TP2	0.63	3.96	-0.18	0.61	4.25	-0.07	0.69	3.37	-0.21	0.65	4.43	-0.50	0.78	3.08	-0.29	0.61	4.80	-0.53	0.78	3.09	-0.30	0.95	1.43	-0.13
CA-TP3	0.68	6.34	0.76	0.70	7.36	1.03	0.73	5.16	0.67	0.72	2.63	0.12	0.78	2.02	0.03	0.71	1.97	-0.02	0.83	1.54	0.03	0.96	0.76	0.03
CA-TP4	0.71	5.89	0.77	0.73	6.77	1.04	0.76	4.88	0.70	0.76	2.39	0.14	0.79	1.96	0.05	0.78	1.83	0.05	0.85	1.52	0.06	0.97	0.72	0.05
CH-Dav	0.46	2.67	-0.21	0.49	2.72	-0.07	0.53	2.82	-0.04	0.54	2.24	-0.36	0.57	2.21	-0.17	0.65	1.82	-0.26	0.77	1.20	-0.03	0.94	0.61	0.00
DE-Hai	0.66	3.54	0.04	0.69	3.85	0.22	0.71	3.10	0.09	0.75	2.65	-0.17	0.67	3.17	-0.27	0.75	3.16	-0.28	0.90	1.61	0.04	0.98	0.70	0.02
DE-Lkb	0.64	3.03	1.09	0.64	3.38	1.39	0.68	3.37	1.40	0.67	1.35	0.32	0.73	1.65	0.68	0.55	2.19	1.36	0.64	1.77	0.92	0.92	0.75	0.44
DE-Lnf	0.66	4.06	0.13	0.68	4.43	0.28	0.74	3.38	0.17	0.76	2.68	-0.17	0.79	2.45	-0.13	0.82	2.50	-0.11	0.91	1.53	0.00	0.98	0.68	0.01
DE-Obe	0.67	3.12	-0.23	0.68	3.10	-0.10	0.75	2.61	-0.15	0.75	3.03	-0.40	0.77	2.91	-0.38	0.66	3.47	-0.36	0.88	1.59	-0.04	0.98	0.72	-0.01
DE-Tha	0.64	3.85	-0.02	0.67	4.02	0.10	0.74	2.97	0.00	0.74	2.80	-0.32	0.76	2.95	-0.37	0.76	3.24	-0.41	0.88	1.51	-0.05	0.98	0.68	-0.01
DK-Sor	0.78	2.94	-0.12	0.77	3.18	0.01	0.81	2.67	-0.04	0.83	3.02	-0.30	0.83	2.91	-0.28	0.84	3.34	-0.28	0.93	1.60	-0.08	0.99	0.74	-0.03
FI-Hyy	0.71	3.56	0.23	0.72	4.15	0.41	0.77	3.35	0.31	0.78	1.92	-0.11	0.85	1.43	-0.13	0.79	1.59	0.01	0.92	1.03	-0.05	0.98	0.47	0.00
FI-Sod	0.68	1.52	-0.24	0.69	1.55	-0.13	0.74	1.37	-0.15	0.74	1.63	-0.49	0.79	1.49	-0.43	0.74	1.42	-0.32	0.87	0.98	-0.13	0.98	0.44	-0.05
FR-LBr	0.48	5.42	0.44	0.54	5.30	0.52	0.60	3.97	0.38	0.60	2.29	-0.19	0.72	1.79	-0.06	0.60	2.04	-0.16	0.72	1.57	0.00	0.95	0.71	0.01
FR-Pue	0.25	4.07	0.35	0.32	8.52	1.76	0.42	5.24	1.01	0.31	3.57	0.50	0.71	1.98	0.24	0.46	1.73	0.21	0.76	0.99	0.09	0.95	0.49	0.06

Sites	CASA			C_Fix			GLO_PEM			VPM			EC_LUE			MODIS			SVM			RF		
	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE	R2	RMSE	RPE
IT-Col	0.74	5.23	0.61	0.74	6.17	0.83	0.79	5.51	0.77	0.81	2.91	0.27	0.76	2.73	0.14	0.63	2.79	0.11	0.90	1.48	-0.07	0.98	0.68	-0.01
IT-Cpz	0.31	5.03	0.22	0.41	10.08	1.27	0.43	6.33	0.71	0.43	4.16	0.33	0.54	1.86	-0.11	0.40	2.04	0.04	0.69	1.34	-0.07	0.94	0.63	-0.03
IT-Lav	0.51	4.09	-0.24	0.50	4.11	-0.11	0.59	3.69	-0.14	0.59	3.85	-0.39	0.55	4.26	-0.47	0.62	3.83	-0.42	0.76	2.15	-0.09	0.95	0.96	-0.03
IT-Ren	0.49	3.34	0.03	0.46	3.75	0.24	0.56	3.49	0.22	0.51	2.73	-0.17	0.53	2.62	-0.20	0.51	2.59	-0.18	0.76	1.83	-0.14	0.96	0.80	-0.04
IT-Ro2	0.62	6.96	1.01	0.65	7.14	1.13	0.65	4.91	0.75	0.68	2.67	0.12	0.79	2.42	0.17	0.60	2.84	0.03	0.87	1.57	0.00	0.98	0.69	0.00
IT-SRo	0.48	6.97	0.73	0.54	7.96	0.95	0.66	4.80	0.56	0.61	2.31	0.04	0.63	1.93	-0.15	0.59	1.78	-0.08	0.76	1.33	-0.08	0.96	0.62	-0.03
NL-Loo	0.72	3.11	0.07	0.74	3.53	0.21	0.80	2.64	0.13	0.81	1.98	-0.28	0.84	1.64	-0.21	0.75	1.96	-0.23	0.90	1.05	0.03	0.98	0.49	0.02
RU-Fyo	0.62	3.54	0.04	0.63	3.88	0.19	0.69	3.05	0.08	0.69	2.52	-0.26	0.74	2.42	-0.32	0.80	2.25	-0.29	0.86	1.67	-0.14	0.97	0.77	-0.06
US-Blo	0.51	7.47	1.13	0.53	8.33	1.37	0.58	6.82	1.16	0.55	2.94	0.35	0.64	3.16	0.43	0.12	2.50	-0.17	0.76	1.19	0.09	0.95	0.59	0.06
US-GBT	0.26	2.83	0.30	0.27	3.04	0.49	0.27	3.15	0.50	0.27	1.98	-0.11	0.28	2.20	0.08	0.33	1.66	-0.23	0.30	1.98	0.23	0.75	0.98	0.07
US-GLE	0.70	1.51	-0.09	0.71	1.62	0.10	0.75	1.61	0.08	0.75	1.43	-0.34	0.79	1.22	-0.23	0.80	1.60	-0.40	0.84	0.98	-0.06	0.96	0.50	-0.02
US-Me2	0.47	5.00	0.31	0.49	5.58	0.54	0.59	4.48	0.38	0.54	2.47	-0.14	0.69	1.98	-0.21	0.72	1.83	-0.24	0.85	1.09	-0.02	0.96	0.58	0.00
US-Me4	0.08	6.78	0.47	0.07	7.38	0.65	0.07	6.54	0.58	0.09	3.74	-0.08	0.09	2.99	-0.18	0.04	3.06	-0.17	0.02	2.39	-0.04	0.34	1.15	-0.03
US-Me6	0.62	0.87	-0.51	0.63	0.61	-0.26	0.66	0.76	-0.35	0.65	0.99	-0.62	0.72	0.74	-0.41	0.51	1.02	-0.60	0.75	0.71	0.35	0.89	0.32	0.14
US-MMS	0.72	7.21	0.89	0.75	6.42	0.90	0.74	4.33	0.53	0.79	2.57	0.10	0.83	2.82	0.23	0.67	3.15	-0.12	0.92	1.45	0.02	0.98	0.68	0.01
US-NR1	0.70	2.96	0.49	0.70	3.69	0.87	0.76	3.57	0.72	0.75	1.66	0.11	0.81	1.96	0.32	0.87	0.99	0.07	0.86	1.13	0.18	0.97	0.55	0.11
US-Oho	0.81	5.99	0.63	0.81	5.71	0.68	0.81	3.90	0.41	0.84	2.23	-0.01	0.91	2.81	0.25	0.77	2.92	-0.14	0.96	1.13	-0.01	0.99	0.55	0.00
US-Prr	0.53	2.56	0.70	0.51	3.00	0.95	0.57	2.75	0.90	0.58	1.37	0.12	0.65	1.25	0.20	0.66	1.41	0.42	0.68	1.26	0.51	0.87	0.65	0.20
US-Syv	0.46	8.18	1.60	0.44	8.65	1.79	0.44	7.53	1.55	0.44	5.54	0.97	0.39	5.62	0.94	0.43	3.50	0.37	0.51	3.07	0.24	0.64	2.62	0.17
US-WCr	0.76	4.23	0.60	0.72	5.18	0.84	0.74	3.96	0.58	0.76	2.52	0.14	0.78	2.76	0.23	0.76	2.38	0.01	0.91	1.43	0.01	0.98	0.70	0.01
ZM-Mon	0.02	7.75	1.34	0.11	7.38	1.37	0.09	4.92	0.82	0.00	3.19	-0.12	0.57	1.84	0.12	0.46	1.99	-0.06	0.80	1.17	0.00	0.97	0.49	0.00
Average	**0.55**	**4.35**	**0.43**	**0.57**	**4.94**	**0.80**	**0.61**	**3.84**	**0.50**	**0.61**	**2.60**	**0.02**	**0.67**	**2.35**	**0.03**	**0.62**	**2.29**	**-0.01**	**0.77**	**1.49**	**0.20**	**0.92**	**0.77**	**0.10**

Table S4. The optimum parameters of RF and SVM across DBF, EBF, ENF, and MF.

Forest types	RF		SVM	
	ntree	mtry	gamma	cost
DBF	250	4	0.4	4.5
EBF	950	2	0.3	9
ENF	400	3	0.6	5
MF	750	1	0.5	9