Table S1: Fluxnet sites used to calculate the weighting. 159 Flunxet sites used to derive DOLCE. The IGBP Land Cover Type Classification at the site level is taken from the site investigator's descriptions found in fluxnet.ornl.gov when available or otherwise from literature citation. Number of site-months indicates the number of site months available in this study for each site. The Dataset is the source of the Data, that is Fluxnet dataset published in Nov, 2016 (FN) or the 2007 LaThuile free fair use dataset (LT).

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
AT-Neu	Neustift/Stubai Valley	Austria	11.3175	47.11667	FN	GRA	96	(Stoy et al., 2013)
AU-AdE	Adelaide River	Australia	131.11780	-13.07690	FN	WSA	19	fluxnet.ornl.gov
AU-DaP	Daly River Pasture	Australia	131.31810	-14.063300	FN	GRA	25	fluxnet.ornl.gov
AU-DaS	Daly River SAV	Australia	131.38800	-14.159283	FN	WSA	24	(Jamali et al., 2011)
AU-Dry	Dry River	Australia	132.37060	-15.258800	FN	WSA	12	fluxnet.ornl.gov
AU-Fog	Fogg Dam	Australia	131.30720	-12.545200	FN	WET	32	(Stoy et al., 2013)
AU-How	Howard Springs	Australia	131.15	-12.4952	LT	SAV	67	(Stoy et al., 2013)
AU-Stp	Sturt Plains	Australia	133.35030	-17.150800	FN	GRA	13	fluxnet.ornl.gov
AU-Tum	Tumbarumba	Australia	148.15160	-35.656600	FN	EBF	102	(Stoy et al., 2013)
AU-Wac	Wallaby Creek	Australia	145.1873	-37.429	LT	EBF	32	(Stoy et al., 2013)
BE-Bra	Brasschaat (De Inslag Forest)	Belgium	4.52056	51.30917	FN	MF	102	(Stoy et al., 2013)
BE-Jal	Jalhay	Belgium	6.07333	50.56389	LT	MF	10	(Stoy et al., 2013)
BE-Lon	Lonzee	Belgium	4.74613	50.55159	FN	CRO	69	(Moureaux, et al., 2006)
BE-Vie	VieIsalm	Belgium	5.99805	50.30507	FN	MF	116	(Stoy et al., 2013)
BR-Sa3	Santarem-Km83-Logged Forest	Brazil	-54.9714	-3.01803	FN	EBF	44	(Stoy et al., 2013)

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
BW-Ma1	Maun- Mopane Woodland	Botswana	23.56033	-19.9165	LT	SAV	19	(Stoy et al., 2013)
CA-Mer	Eastern Peatland-Mer Bleue	Canada	-75.5186	45.4094	LT	WET	70	(Lafleur, et al., 2003)
CA-Qcu	Quebec - Eastern Boreal, Black Spruce/Jack Pine Cutover	Canada	-74.0365	49.2671	LT	ENF	64	fluxnet.ornl.gov
CA-Qfo	QC-Eastern Old Black Spruce (EOBS)	Canada	-74.3421	49.6925	FN	ENF	74	(Bergeron et al., 2007)
CA-SF1	Saskatchewan - Western Boreal, forest burned in 1977	Canada	-105.8176	54.4850	FN	ENF	40	(Mkhabela et al., 2009)
CA-SF2	Saskatchewan - Western Boreal, forest burned in 1989	Canada	-105.8775	54.2539	FN	ENF	48	(Mkhabela et al., 2009)
CH-Cha	Chamau	Switzerland	8.41044	47.210222	FN	GRA	44	(Merbold et al., 2014)
CH-Dav	Davos- Seehorn forest	Switzerland	9.8559	46.8153	FN	ENF	113	fluxnet.ornl.gov
CH-Fru	Früebüel	Switzerland	8.53778	47.115833	FN	GRA	45	(Imer et al., 2013)
CN-Cng	Changling	China	123.5092	44.5934	FN	GRA	31	fluxnet.ornl.gov
CN-Du2	Duolun_GRA	China	116.28360	42.046700	FN	GRA	25	(Stoy et al., 2013)
CN-HaM	Haibei Alpine Tibet site	China	101.18	37.37	FN	GRA	36	(Li et al., 2013)
CZ-wet	CZECHWET	Czech Republic	14.77035	49.024650	LT	WET	9	(Stoy et al., 2013)

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
DE-Geb	Gebesee	Germany	10.9143	51.1001	LT	CRO	107	(Revill et al., 2013)
DE-Gri	Grillenburg- grass station	Germany	13.51253	50.94947	FN	GRA	71	(Stoy et al., 2013)
DE-Hai	Hainich	Germany	10.453	51.07917	FN	DBF	120	(Ershadi et al., 2014)
DE-Kli	Klingenberg - CRO	Germany	13.52251	50.89288	FN	CRO	60	(Revill et al., 2013)
DE-Lkb	Lackenberg	Germany	13.3047	49.0996	FN	ENF	7	fluxnet.ornl.gov
DE-Meh	Mehrstedt 1	Germany	10.65547	51.27531	LT	MF	40	fluxnet.ornl.gov
DE-Obe	Oberbärenburg	Germany	13.7196	50.7836	FN	ENF	23	(Jung et al., 2009)
DE-Seh	Selhausen	Germany	6.4496531	50.870623	FN	CRO	31	fluxnet.ornl.gov
DE-Tha	Anchor Station Tharandt - old spruce	Germany	13.56694	50.96361	FN	ENF	120	(Wei et al., 2014)
DE-Wet	Wetzstein	Germany	11.45753	50.4535	LT	ENF	60	(Stoy et al., 2013)
DK-Fou	Foulum	Denmark	9.58722	56.4842	LT	CRO	8	(Stoy et al., 2013)
DK-Lva	Lille Valby (Rimi)	Denmark	12.08330	55.683300	LT	GRA	23	(Stoy et al., 2013)
DK-Ris	Risbyholm	Denmark	12.09722	55.530278	LT	CRO	24	(Gilmanov et al., 2010)
DK-Sor	Soroe	Denmark	11.64464	55.485869	FN	DBF	102	(Stoy et al., 2013)
DK-ZaF	Zackenberg Fen	Denmark	-20.5545	74.4814	FN	WET	3	(Soegaard and Nordstroem, 1999)
ES-ES1	El Saler	Spain	-0.31881	39.34597	LT	ENF	78	(Stoy et al.,

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
								2013)
ES-ES2	El Saler-Sueca	Spain	-0.31528	39.27556	LT	CRO	32	(Stoy et al., 2013)
ES-LgS	Laguna Seca	Spain	-2.9658	37.0979	FN	OSH	30	fluxnet.ornl.gov
ES-LMa	Las Majadas del Tietar	Spain	-5.77336	39.9415	LT	SAV	36	(Stoy et al., 2013)
ES-VDA	Vall d'Alinya	Spain	1.4485	42.15218	LT	GRA	28	(Stoy et al., 2013)
FI-Hyy	Hyytiala	Finland	24.2950	61.8475	FN	ENF	80	(Stoy et al., 2013)
FI-jok	Jokioinen	Finland	23.5135	60.8986	FN	CRO	39	(Reichstein et al., 2005)
FI-Kaa	Kaamanen wetland	Finland	27.2950	69.1407	FN	WET	82	(Aurela et al., 2001)
FI-Lom	Lompolojänkkä	Finland	24.2092	67.9972	FN	WET	36	fluxnet.ornl.gov
FI-Sod	Sodankyla	Finland	26.6378	67.3619	FN	ENF	110	(Stoy et al., 2013)
FR-Gri	Grignon	France	1.95191	48.84422	FN	CRO	68	(Loubet et al., 2011)
FR-Hes	Hesse Forest-Sarrebourg	France	7.06556	48.67416	LT	DBF	80	(Reichstein et al., 2005)
FR-LBr	Le Bray (after 6/28/1998)	France	-0.7693	44.71711	LT	ENF	102	(Stoy et al., 2013)
FR-Lq1	Laqueuille	France	2.73583	45.64306	LT	GRA	36	(Gilmanov et al., 2007)
FR-Lq2	Laqueuille extensive	France	2.73703	45.63919	LT	GRA	36	(Gilmanov et al., 2007)
FR-Pue	Puechabon	France	3.59583	43.741390	FN	EBF	83	(Wei et al., 2014)
HU-Bug	Bugacpuszta	Hungary	19.6013	46.6911	LT	GRA	50	(Stoy et al., 2013)
HU-Mat	Matra	Hungary	19.726	47.8469	LT	GRA	32	(Stoy et al.,

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
								2013)
ID-Pag	Palangkaraya (PDF)	Indonesia	114.036	-2.345	LT	EBF	18	(Hirano et al., 2007)
IE-Ca1	Carlow1	Ireland	-6.9181	52.8588	LT	CRO	33	(Stoy et al., 2013)
IE-Dri	Dripsey	Ireland	-8.75181	51.986694	LT	GRA	31	(Stoy et al., 2013)
IL-Yat	Yatir	Israel	35.0515	31.3450	FN	ENF	67	(Reichstein et al., 2003)
IT-Amp	Amplero	Italy	13.60516	41.9041	LT	GRA	54	(Gilmanov et al., 2007)
IT-Cas	Castellaro	Italy	8.717522222	45.07004722	LT	CRO	6	(Stoy et al., 2013)
IT-Col	Collelongo- Selva Piana	Italy	13.58814	41.84936	LT	DBF	72	(Stoy et al., 2013)
IT-Cpz	Castelporziano	Italy	12.3761	41.7052	FN	EBF	94	(Wei et al., 2014)
IT-Lav	Le Bray (after 6/28/1998)	Italy	11.28132	45.9562	FN LT	ENF	108	(Cescatti & Zorer, 2003)
IT-Lec	Lecceto	Italy	11.26975	43.30359	LT	EBF	17	(Stoy et al., 2013)
IT-LMa	La Mandria	Italy	7.58259	45.15258	LT	DBF	30	fluxnet.ornl.gov
IT-Mal	Malga Arpaco	Italy	11.70334	46.11402	LT	GRA	32	(Gilmanov et al., 2007)
IT-MBo	Monte Bondone	Italy	11.04583	46.01468	LT	GRA	84	(Gilmanov et al., 2007)
IT-Non	Nonantola	Italy	11.09109	44.69019	LT	MF	40	fluxnet.ornl.gov
IT-PT1	Zerbolo-Parco Ticino- Canarazzo	Italy	9.06104	45.20087	FN	DBF	31	(Stoy et al., 2013)
IT-Ren	Renon/Ritten (Bolzano)	Italy	11.43369	46.58686	FN	ENF	80	(Stoy et al., 2013)
IT-Ro2	Roccarespampani 2	Italy	11.92093	42.390260	FN	DBF	77	(Tedeschi et

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
								al., 2006)
IT-SRo	San Rossore	Italy	10.28444	43.727861	FN	ENF	109	(Reichstein et al., 2005)
IT-Tor	Torgnon	Italy	7.5781	45.8444	FN	GRA	16	(Galvagno et al., 2013)
NL-Ca1	Cabauw	Netherland	4.927	51.971	LT	GRA	46	fluxnet.ornl.gov
NL-Haa	Haastrecht	Netherland	4.80556	52.003611	LT	GRA	10	fluxnet.ornl.gov
NL-Hor	Horstermeer	Netherland	5.07130	52.240350	FN	GRA	50	(Sulkava et al., 2011)
NL-Lan	Langerak	Netherland	4.9029	51.9536	LT	CRO	11	(Stoy et al., 2013)
NL-Loo	Loobos	Netherland	5.74356	52.166581	FN	ENF	119	(Gash and Dolman, 2003)
NL-Lut	Lutjewad	Netherland	6.35603	53.398922	LT	CRO	9	(Stoy et al., 2013)
NL-Mol	Molenweg	Netherland	4.63908	51.65	LT	CRO	7	(Gilmanov et al., 2007)
PL-Wet	Polwet	Poland	16.3094	52.7622	LT	WET	24	fluxnet.ornl.gov
PT-Esp	Espirra	Portugal	-8.6018	38.6394	LT	EBF	43	(Wei et al., 2014)
PT-Mi1	Mitra (Evora)	Portugal	-8.00006	38.54064	LT	EBF	20	(Reichstein et al., 2003)
PT-Mi2	Mitra IV Tojal	Portugal	-8.02455	38.4765	LT	GRA	29	(Gilmanov et al., 2007)
RU-Cok	Chokurdakh	Russia	147.4943	70.8291	FN	OSH	29	(Stoy et al., 2013)
RU-Fyo	Fyodorovskoye wet spruce stand	Russia	32.92208	56.46153	FN	ENF	118	(Stoy et al., 2013)
RU-Ha1	Hakasia steppe	Russia	90.00215	54.72517	FN	GRA	19	(Belelli Marchesini

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
								et al., 2007)
RU-Ha2	Ubs Nur- Hakasija-recovering grassland	Russia	89.95664	54.77301	LT	GRA	8	(Belelli Marchesini et al., 2007)
RU-Ha3	Ubs Nur-Hakasija-Site 3	Russia	89.07785	54.70455	LT	GRA	7	(Belelli Marchesini et al., 2007)
RU-Zot	Zotino	Russia	89.3508	60.8008	FN	ENF	19	(Eugster et al., 2000)
SE-Deg	Degero	Sweden	19.5500	64.1833	FN	GRA	56	fluxnet.ornl.gov
SE-Faj	Fajemyr	Sweden	13.55351	56.26547	LT	WET	12	(Stoy et al., 2013)
SE-Fla	Flakaliden	Sweden	19.4569	64.1128	FN	ENF	35	(Stoy et al., 2013)
SE-Nor	Norunda	Sweden	17.4795	60.0865	FN	ENF	24	(Stoy et al., 2013)
SE-Sk1	Skyttorp1 young	Sweden	17.9181	60.1250	FN	ENF	11	(Stoy et al., 2013)
SE-Sk2	Skyttorp	Sweden	17.8401	60.1297	FN	ENF	15	(Stoy et al., 2013)
SK-Tat	Tatra	Slovak Republic	20.1635	49.12078	LT	ENF	4	(Stoy et al., 2013)
UK-AMo	Auchencorth Moss-Scotland	UK	-3.23889	55.79167	LT	GRA	8	fluxnet.ornl.gov
UK-EBu	Easter Bush-Scotland	UK	-3.20578	55.866	LT	GRA	16	(Stoy et al., 2013)
UK-ESa	East Saltoun	UK	-2.85861	55.90694	LT	CRO	22	(Stoy et al., 2013)
UK-Gri	Griffin- Aberfeldy-Scotland	UK	-3.7981	56.6072	FN	ENF	35	(Stoy et al., 2013)
UK-Ham	Hampshire	UK	-0.8583	51.15353	LT	DBF	22	(Stoy et al., 2013)
UK-Her	Hertfordshire	UK	-0.476080	51.7837982	LT	CRO	6	(Stoy et al.,

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
				177				2013)
UK-PL3	Pang/ Lambourne (forest)	UK	-1.26667	51.45	LT	DBF	24	fluxnet.ornl.gov
UK-Tad	Tadham Moor	UK	-2.82864	51.2071	LT	GRA	9	fluxnet.ornl.gov
US-AR1	ARM USDA UNL OSU Woodward Switchgrass 1	USA	-99.4200	36.4267	FN	GRA	8	fluxnet.ornl.gov
US-AR2	ARM USDA UNL OSU Woodward Switchgrass 2	USA	-99.5975	36.6358	FN	GRA	7	fluxnet.ornl.gov
US-ARb	ARM Southern Great Plains burn site- Lamont	USA	-98.0402	35.5497	FN	GRA	20	(Stoy et al., 2013)
US-ARc	ARM Southern Great Plains control site- Lamont	USA	-98.0400	35.5465	FN	GRA	20	(Stoy et al., 2013)
US-ARM	OK - ARM Southern Great Plains site- Lamont	USA	-97.4888	36.6058	FN	CRO	84	(Stoy et al., 2013)
US-Aud	AZ - Audubon Research Ranch	USA	-110.51	31.5907	LT	GRA	49	(Hilton et al., 2014)
US-Bar	NH - Bartlett Experimental Forest	USA	-71.2881	44.0646	LT	MF	23	fluxnet.ornl.gov
US-Bkg	Brookings	USA	-96.8362	44.3453	FN	GRA	31	(Gilmanov et al., 2005)
US-Blo	CA - Blodgett Forest	USA	-120.633	38.8952	FN	ENF	91	(Reichstein et al., 2003)
US-Bo1	IL - Bondville	USA	-88.2904	40.0062	LT	CRO	83	(Stoy et al., 2013)
US-Cop	Corral Pocket	USA	-109.3900	38.0900	FN	GRA	48	fluxnet.ornl.gov
US-FPe	MT - Fort Peck	USA	-105.102	48.3077	LT	GRA	75	(Ershadi et al., 2014)

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
US-GBT	GLEES Brooklyn Tower	USA	-106.2397	41.3658	FN	ENF	63	(Zeller and Hehn, 1996)
US-GLE	GLEES	USA	-106.2399	41.3665	FN	ENF	57	(Frank et al., 2014)
US-Goo	MS - Goodwin Creek	USA	-89.8735	34.2547	LT	VEG	48	fluxnet.ornl.gov
US-Ho1	ME - Howland Forest (main tower)	USA	-68.7402	45.2041	LT	MF	60	(Wei et al., 2014)
US-Ho2	ME - Howland Forest (west tower)	USA	-68.747	45.2091	LT	MF	60	(Wei et al., 2014)
US-Los	WI - Lost Creek	USA	-89.9792	46.0827	FN	CSH	98	(Sulman et al., 2010)
US-Me1	Metolius - Eyerly burn	USA	-121.5000	44.5794	FN	ENF	10	(Irvine & Hibbard, 2007)
US-Me2	Metolius-intermediate aged ponderosa pine	USA	-121.5574	44.4523	FN	ENF	89	(Irvine & Hibbard, 2007)
US-Me4	Metolius-old aged ponderosa pine	USA	-121.6224	44.4992	FN	ENF	12	(Irvine & Hibbard, 2007)
US-MMS	IN - Morgan Monroe State Forest	USA	-86.4131	39.3231	FN	DBF	120	(Baldocchi et al., 2001)
US-MOz	MO - Missouri Ozark Site	USA	-92.2	38.7441	LT	DBF	31	(Stoy et al., 2013)
US-Ne3	NE - Mead - rainfed maize-soybean rotation site	USA	-96.4397	41.1797	FN	CRO	103	(Verma et al., 2005)
US-NR1	Niwot Ridge Forest (LTER NWT1)	USA	-105.54600	40.032902	FN	ENF	120	(Monson et al., 2002)
US-Oho	OH - Oak Openings	USA	-83.8438	41.5545	LT	DBF	24	(Stoy et al., 2013)
US-SP1	FL - Slashpine-Austin Cary- 65yrs	USA	-82.2188	29.7381	LT	ENF	16	(Stoy et al., 2013)

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
	nat regen							
US-SP2	FL - Slashpine-Mize-clearcut-3yr,regen	USA	-82.2448	29.7648	LT	ENF	60	(Hilton et al., 2014)
US-SP3	FL - Slashpine-Donaldson-mid-rot-12yrs	USA	-82.1633	29.7548	LT	ENF	54	(Hilton et al., 2014)
US-SRG	Santa Rita GRA	USA	-110.8277	31.7894	FN	GRA	21	(Scott et al., 2010)
US-SRM	AZ - Santa Rita Mesquite	USA	-110.866	31.8214	FN	WSA	72	(Scott et al., 2010)
US-Syv	MI - Sylvania Wilderness Area	USA	-89.3477	46.242	FN	MF	65	(Desai, et al., 2005)
US-Ton	CA - Tonzi Ranch	USA	-120.966	38.4316	FN	DBF	102	(Vargas et al., 2011)
US-Var	CA - Vaira Ranch-Ione	USA	-120.951	38.4133	FN	GRA	109	(Ma et al., 2007)
US-WCr	WI - Willow Creek	USA	-90.0799	45.8059	FN	DBF	79	(Cook et al., 2004)
US-Whs	Walnut Gulch Lucky Hills Shrub	USA	-110.0522	31.7438	FN	OSH	30	(Cavanaugh, et al., 2011)
US-Wi0	WI - Young red pine (YRP)	USA	-91.0814	46.6188	LT	ENF	6	(Noormets et al., 2007)
US-Wi1	WI - Intermediate hardwood	USA	-91.2329	46.7305	LT	DBF	5	(Noormets et al., 2007)
US-Wi3	Mature hardwood	USA	-91.0987	46.6347	FN	DBF	16	(Noormets et al., 2007)
US-Wi4	WI - Mature red pine	USA	-91.1663	46.7393	LT	ENF	23	(Hilton et al., 2014)
US-Wi5	WI - Mixed young jack pine	USA	-91.0858	46.6531	LT	ENF	7	(Stoy et al., 2013)
US-Wi6	WI - Pine barrens #1	USA	-91.2982	46.6249	LT	OSH	11	(Noormets et al., 2007)
US-Wi7	WI - Red pine	USA	-91.0693	46.6491	LT	MF	5	fluxnet.ornl.

Site code	Site name	Country	Longitude	Latitude	Dataset	IGBP at the site level	Number of site-months	Reference
	clearcut							gov
US-Wi8	WI - Young hardwood clearcut	USA	-91.2524	46.7223	LT	DBF	5	(Noormets et al., 2007)
US-Wkg	AZ - Walnut Gulch Kendall grasslands	USA	-109.942	31.7365	FN	GRA	68	(Scott et al., 2010)
ZA-Kru	Skukuza- Kruger National Park	South Africa	31.4969	-25.0197	FN	SAV	69	(King et al., 2003)
ZM-Mon	Mongu	Zambia	23.2528	-15.4378	FN	DBF	22	(King et al., 2003)

In-sample test

The in-sample test shows that the weighting approach is monotonically not increasing (MSE) and that it converges to the same results no matter where we start. This confirms that the weighting technique always builds a better performing ET estimates from an ensemble of ET products relative to the in-sample observational dataset (Fig. S1).

Figure S1: The results of the in-sample test showing how the Mean Square Error (MSE) of the weighting changes when we increase the number of ET products involved in the weighting from 1 to 6. The test was repeated 100 times of a random selection of n products ($2 \leq n \leq 6$).

Clustering test structure and results

We assigned one of the five following biome groups to each flux tower site. Each biome group is a composite of two or more land cover types. We performed the combination based on the similarity of the combined components, by doing that we increase the available number of sites in each cluster:

- MF DBF SH: sites covered by mixed forest (MF), deciduous broadleaf forest (DBF) or Shrublands (Open (OSH) or closed (CSH))
- ENF EBF: includes evergreen needleleaf (ENF) and evergreen broadleaf forest (EBF) sites
- SAV WSA GRA: includes sites covered by Savannas (SAV),woody Savannas (WSA) or Grasslands
- CRO VEG: includes croplands (CRO) and grasslands (GRA)
- Remaining: includes the the Sites that don't belong to any of the four groups defined above were put together in a "remaining group".

In each particular cluster, one site was taken out and the remaining sites were used to train the weighting and derive a set of weights and bias correction terms. The mean of the weighting ensemble was calculated and the derived cluster-dependent weighted product was evaluated against the out-sample site in a particular cluster. This has resulted in calculating the statistical metrics MSE, RSD and COR – described in Sec. 2.4 separately at every cluster. In a parallel analysis, a non-clustered weighting was performed as explained in Sec. 2.4 The weights generated in the non-clustered weighting were assigned to the participating products at the sites belonging to the remaining group in the clustered weighting. The weighting was repeated for all the sites being taken out of sample.

In the clustering tests we looked at:

1. The performance improvement (%) of the cluster-independent weighted product against the mean of the diagnostic ensemble (No CL).
2. The performance improvement (%) of the cluster-independent weighted product against each individual dataset in Ref, this yielded 10 different values of performance improvement. Then, these values were aggregated in one boxplot (No CL Agg).
3. The performance improvement (%) of the cluster-dependent weighted product against the mean of the diagnostic ensemble (CL-5).
4. The performance improvement (%) of the cluster-dependent weighted product against each individual dataset in Ref, this yielded 10 different values of performance improvement. Then, these values were aggregated in one boxplot (CL-5 Agg).
5. The performance improvement (%) of the cluster-dependent weighted product against the mean of the diagnostic ensemble at a particular cluster. This was calculated for all the clusters.
6. The performance improvement (%) of the cluster-dependent weighted product against each individual dataset in Ref at a particular cluster, this yielded 10 different values of performance improvement for each cluster, which were then aggregated and displayed as a single boxplot for every cluster.

The results of the clustered test in Fig. S2 show a significant difference in the performance across the individual biome groups. The median value of the cluster-dependent weighting has slightly decreased which indicates that the biome types classification doesn't help to better align to the flux behaviour.

Figure S2*: Box and whisker plots displaying the results of the* One out-of-sample site test for the cluster independent weighting (No CL and No CL Agg boxplots), the cluster dependent weighting (Cl-5 and CL-5 Agg) and over individual biome types for three metrics: MSE (a), RSD (b) and COR (c).

References

Anon: Home | fluxnetweb.ornl.gov, [online] Available from: https://fluxnet.ornl.gov/ (Accessed 1 February 2016), n.d.

Aurela, M., Laurila, T. and Tuovinen, J.-P.: Seasonal CO2 balances of a subarctic mire, J. Geophys. Res. Atmos., 106(D2), 1623–1637, doi:10.1029/2000JD900481, 2001.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T.,

Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bull. Am. Meteorol. Soc., 82(11), 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Belelli Marchesini, L., Papale, D., Reichstein, M., Vuichard, N., Tchebakova, N. and Valentini, R.: Carbon balance assessment of a natural steppe of southern Siberia by multiple constraint approach, Biogeosciences, 4, 581–595 [online] Available from: www.biogeosciences.net/4/581/2007/ (Accessed 13 March 2017), 2007.

Bergeron, O., Margolis, H. A., Black, T. A., Coursolle, C., Dunn, A. L., Barr, A. G. and Wofsy, S. C.: Comparison of carbon dioxide fluxes over three boreal black spruce forests in Canada, Glob. Chang. Biol., 13(1), 89–107, doi:10.1111/j.1365-2486.2006.01281.x, 2007.

Cavanaugh, M. L., Kurc, S. A. and Scott, R. L.: Evapotranspiration partitioning in semiarid shrubland ecosystems: a two-site evaluation of soil moisture control on transpiration, Ecohydrology, 4(5), 671–681, doi:10.1002/eco.157, 2011.

CESCATTI, A. and ZORER, R.: Structural acclimation and radiation regime of silver fir (Abies alba Mill.) shoots along a light gradient, Plant, Cell Environ., 26(3), 429–442, doi:10.1046/j.1365-3040.2003.00974.x, 2003.

Cook, B. D., Davis, K. J., Wang, W., Desai, A., Berger, B. W., Teclaw, R. M., Martin, J. G., Bolstad, P. V., Bakwin, P. S., Yi, C. and Heilman, W.: Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA, Agric. For. Meteorol., 126(3–4), 271–295, doi:10.1016/j.agrformet.2004.06.008, 2004.

Desai, A. R., Bolstad, P. V., Cook, B. D., Davis, K. J. and Carey, E. V.: Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA, Agric. For. Meteorol., 128(1–2), 33–55, doi:10.1016/j.agrformet.2004.09.005, 2005.

Ershadi, A., McCabe, M. F., Evans, J. P., Chaney, N. W. and Wood, E. F.: Multi-site evaluation of terrestrial evaporation models using FLUXNET data, Agric. For. Meteorol., doi:10.1016/j.agrformet.2013.11.008, 2014.

Eugster, W., Rouse, W. R., Pielke Sr, R. A., Mcfadden, J. P., Baldocchi, D. D., Kittel, T. G. F., Chapin, F. S., Liston, G. E., Vidale, P. L., Vaganov, E. and Chambers, S.: Land-atmosphere energy exchange in Arctic tundra and boreal forest: available data and feedbacks to climate, Glob. Chang. Biol., 6(S1), 84–115, doi:10.1046/j.1365-2486.2000.06015.x, 2000.

Frank, J. M., Massman, W. J., Ewers, B. E., Huckaby, L. S., Negron, J. F., Frank, J. M., Massman, W. J., Ewers, B. E., Huckaby, L. S. and Negron, J. F.: Ecosystem CO2/H2O fluxes are explained by hydraulically limited gas exchange during tree mortality from spruce bark beetles, J. Geophys. Res. Biogeosciences, 119(6), 1195–1215 [online] Available from: https://www.treesearch.fs.fed.us/pubs/46894 (Accessed 13 March 2017), 2014.

Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., Julitta, T., Manca, G., Siniscalco, C., Morra di Cella, U. and Migliavacca, M.: Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season, Environ. Res. Lett., 8(2), 25008, doi:10.1088/1748-9326/8/2/025008, 2013.

Gash, J. H. C. and Dolman, A. J.: Sonic anemometer (co)sine response and flux measurement I. The potential for (co)sine error to affect sonic anemometer-based flux measurements, Agric. For. Meteorol., 119, 195–207, doi:10.1016/S0168-1923(03)00137-0, 2003.

Gilmanov, T. G., Tieszen, L. L., Wylie, B. K., Flanagan, L. B., Frank, A. B., Haferkamp, M. R., Meyers, T. P. and Morgan, J. A.: Integration of CO_2 flux and remotely-sensed data for primary production and ecosystem respiration analyses in the Northern Great Plains: potential for quantitative spatial extrapolation, Glob. Ecol. Biogeogr., 14(3), 271–292, doi:10.1111/j.1466-822X.2005.00151.x, 2005.

Gilmanov, T. G., Soussana, J. F., Aires, L., Allard, V., Ammann, C., Balzarolo, M., Barcza, Z., Bernhofer, C., Campbell, C. L., Cernusca, A., Cescatti, A., Clifton-Brown, J., Dirks, B. O. M., Dore, S., Eugster, W., Fuhrer, J., Gimeno, C., Gruenwald, T., Haszpra, L., Hensen, A., Ibrom, A., Jacobs, A. F. G., Jones, M. B., Lanigan, G., Laurila, T., Lohila, A., G.Manca, Marcolla, B., Nagy, Z., Pilegaard, K., Pinter, K., Pio, C., Raschi, A., Rogiers, N., Sanz, M. J., Stefani, P., Sutton, M., Tuba, Z., Valentini, R., Williams, M. L. and Wohlfahrt, G.: Partitioning European grassland net ecosystem CO2 exchange into gross primary productivity and ecosystem respiration using light response function analysis, Agric. Ecosyst. Environ., 121(1–2), 93–120, doi:10.1016/j.agee.2006.12.008, 2007.

Gilmanov, T. G., Aires, L., Barcza, Z., Baron, V. S., Belelli, L., Beringer, J., Billesbach, D., Bonal, D., Bradford, J., Ceschia, E., Cook, D., Corradi, C., Frank, A., Gianelle, D., Gimeno, C., Gruenwald, T., Guo, H., Hanan, N., Haszpra, L., Heilman, J., Jacobs, A., Jones, M. B., Johnson, D. A., Kiely, G., Li, S., Magliulo, V., Moors, E., Nagy, Z., Nasyrov, M., Owensby, C., Pinter, K., Pio, C., Reichstein, M., Sanz, M. J., Scott, R., Soussana, J. F., Stoy, P. C., Svejcar, T., Tuba, Z. and Zhou, G.: Productivity, Respiration, and Light-Response Parameters of World Grassland and Agroecosystems Derived From Flux-Tower Measurements, Rangel. Ecol. Manag., 63(1), 16–39, doi:10.2111/REM-D-09-00072.1, 2010.

Hilton, T. W., Davis, K. J. and Keller, K.: Evaluating terrestrial CO2 flux diagnoses and uncertainties from a simple land surface model and its residuals, Biogeosciences, doi:10.5194/bg-11-217-2014, 2014.

HIRANO, T., SEGAH, H., HARADA, T., LIMIN, S., JUNE, T., HIRATA, R. and OSAKI, M.: Carbon dioxide balance of a tropical peat swamp forest in Kalimantan, Indonesia, Glob. Chang. Biol., 13(2), 412–425, doi:10.1111/j.1365-2486.2006.01301.x, 2007.

Imer, D., Merbold, L., Eugster, W. and Buchmann, N.: Temporal and spatial variations of soil CO 2 , CH 4 and N 2 O fluxes at three differently managed grasslands, Biogeosciences, 10, 5931–5945, doi:10.5194/bg-10-5931-2013, 2013.

IRVINE, J., LAW, B. E. and HIBBARD, K. A.: Postfire carbon pools and fluxes in semiarid ponderosa pine in Central Oregon, Glob. Chang. Biol., 13(8), 1748–1760, doi:10.1111/j.1365-2486.2007.01368.x, 2007.

Jamali, H., Livesley, S. J., Dawes, T. Z., Cook, G. D., Hutley, L. B. and Arndt, S. K.: Diurnal and seasonal variations in CH4 flux from termite mounds in tropical savannas of the Northern Territory, Australia, Agric. For. Meteorol., 151(11), 1471–1479, doi:10.1016/j.agrformet.2010.06.009, 2011.

Jung, M., Reichstein, M. and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6(10), 2001–2013, doi:10.5194/bg-

6-2001-2009, 2009.

King, M. D., Platnick, S., Moeller, C. C., Revercomb, H. E. and Chu, D. A.: Remote sensing of smoke, land, and clouds from the NASA ER-2 during SAFARI 2000, J. Geophys. Res. Atmos., 108(D13), n/a-n/a, doi:10.1029/2002JD003207, 2003.

Lafleur, P. M., Roulet, N. T., Bubier, J. L., Frolking, S. and Moore, T. R.: Interannual variability in the peatland-atmosphere carbon dioxide exchange at an ombrotrophic bog, Global Biogeochem. Cycles, 17(2), n/a-n/a, doi:10.1029/2002GB001983, 2003.

Li, X., Liang, S., Yu, G., Yuan, W., Cheng, X., Xia, J., Zhao, T., Feng, J., Ma, Z., Ma, M., Liu, S., Chen, J., Shao, C., Li, S., Zhang, X., Zhang, Z., Chen, S., Ohta, T., Varlagin, A., Miyata, A., Takagi, K., Saiqusa, N. and Kato, T.: Estimation of gross primary production over the terrestrial ecosystems in China, Ecol. Modell., 261–262, 80–92, doi:10.1016/j.ecolmodel.2013.03.024, 2013.

Loubet, B., Laville, P., Lehuger, S., Larmanou, E., Fléchard, C., Mascher, N., Genermont, S., Roche, R., Ferrara, R. M., Stella, P., Personne, E., Durand, B., Decuq, C., Flura, D., Masson, S., Fanucci, O., Rampon, J.-N., Siemens, J., Kindler, R., Gabrielle, B., Schrumpf, M. and Cellier, P.: Carbon, nitrogen and Greenhouse gases budgets over a four years crop rotation in northern France, Plant Soil, 343(1–2), 109–137, doi:10.1007/s11104-011-0751-9, 2011.

Ma, S., Baldocchi, D. D., Xu, L. and Hehn, T.: Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California, Agric. For. Meteorol., doi:10.1016/j.agrformet.2007.07.008, 2007.

Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D. and Buchmann, N.: Greenhouse gas budget CO2, CH4 and N2O) of intensively managed grassland following restoration, Glob. Chang. Biol., 20(6), 1913–1928, doi:10.1111/gcb.12518, 2014.

Mkhabela, M. S., Amiro, B. D., Barr, A. G., Black, T. A., Hawthorne, I., Kidston, J., McCaughey, J. H., Orchansky, A. L., Nesic, Z., Sass, A., Shashkov, A. and Zha, T.: Comparison of carbon dynamics and water use efficiency following fire and harvesting in Canadian boreal forests, Agric. For. Meteorol., 149(5), 783–794, doi:10.1016/j.agrformet.2008.10.025, 2009.

Monson, R. K., Turnipseed, A. A., Sparks, J. P., Harley, P. C., Scott-Denton, L. E., Sparks, K. and Huxman, T. E.: Carbon sequestration in a high-elevation, subalpine forest, Glob. Chang. Biol., 8(5), 459–478, doi:10.1046/j.1365-2486.2002.00480.x, 2002.

Moureaux, C., Debacq, A., Bodson, B., Heinesch, B. and Aubinet, M.: Annual net ecosystem carbon exchange by a sugar beet crop, Agric. For. Meteorol., 139(1–2), 25–39, doi:10.1016/j.agrformet.2006.05.009, 2006.

Noormets, A., Chen, J. and Crow, T. R.: Age-Dependent Changes in Ecosystem Carbon Fluxes in Managed Forests in Northern Wisconsin, USA, Ecosystems, 10(2), 187–203, doi:10.1007/s10021-007-9018-y, 2007.

Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y., Grünzweig, J. M., Irvine, J., Joffre, R., Law, B. E., Loustau, D., Miglietta, F., Oechel, W., Ourcival, J.-M., Pereira, J. S., Peressotti, A., Ponti, F., Qi, Y., Rambal, S., Rayment, M., Romanya, J., Rossi, F., Tedeschi, V., Tirone, G., Xu, M. and Yakir, D.: Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, Global Biogeochem. Cycles, 17(4), n/a-n/a, doi:10.1029/2003GB002035, 2003.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Glob. Chang. Biol., 11(9), 1424–1439, doi:10.1111/j.1365-2486.2005.001002.x, 2005.

Revill, A., Sus, O., Barrett, B. and Williams, M.: Carbon cycling of European croplands: A framework for the assimilation of optical and microwave Earth observation data, Remote Sens. Environ., 137, 84–93, doi:10.1016/j.rse.2013.06.002, 2013.

Scott, R. L., Hamerlynck, E. P., Jenerette, G. D., Moran, M. S. and Barron-gafford, G. A.: Carbon dioxide exchange in a semidesert grassland through drought-induced vegetation change, J. Geophys. Res, 115, doi:10.1029/2010JG001348, 2010.

Soegaard, H. and Nordstroem, C.: Carbon dioxide exchange in a high-arctic fen estimated by eddy covariance measurements and modelling, Glob. Chang. Biol., 5(5), 547–562, doi:10.1111/j.1365-2486.1999.00250.x, 1999.

Stoy, P. C., Mauder, M., Foken, T., Marcolla, B., Boegh, E., Ibrom, A., Arain, M. A., Arneth, A., Aurela, M., Bernhofer, C., Cescatti, A., Dellwik, E., Duce, P., Gianelle, D., van Gorsel, E., Kiely, G., Knohl, A., Margolis, H., Mccaughey, H., Merbold, L., Montagnani, L., Papale, D., Reichstein, M., Saunders, M., Serrano-Ortiz, P., Sottocornola, M., Spano, D., Vaccari, F. and Varlagin, A.: A data-driven analysis of energy balance closure across FLUXNET research sites: The role of landscape scale heterogeneity, Agric. For. Meteorol., 171–172, 137–152, doi:10.1016/j.agrformet.2012.11.004, 2013.

Sulkava, M., Luyssaert, S., Zaehle, S. and Papale, D.: Assessing and improving the representativeness of monitoring networks: The European flux tower network example, J. Geophys. Res., 116(G3), G00J04, doi:10.1029/2010JG001562, 2011.

Sulman, B. N., Desai, A. R., Saliendra, N. Z., Lafleur, P. M., Flanagan, L. B., Sonnentag, O., MacKay, D. S., Barr, A. G. and Van Der Kamp, G.: CO2 fluxes at northern fens and bogs have opposite responses to inter-annual fluctuations in water table, Geophys. Res. Lett., doi:10.1029/2010GL044018, 2010.

TEDESCHI, V., REY, A., MANCA, G., VALENTINI, R., JARVIS, P. G. and BORGHETTI, M.: Soil respiration in a Mediterranean oak forest at different developmental stages after coppicing, Glob. Chang. Biol., 12(1), 110–121, doi:10.1111/j.1365-2486.2005.01081.x, 2006.

Vargas, R., Baldocchi, D. D., Bahn, M., Hanson, P. J., Hosman, K. P., Kulmala, L., Pumpanen, J. and Yang, B.: On the multi-temporal correlation between photosynthesis and soil CO 2 efflux: Reconciling lags and observations, New Phytol., doi:10.1111/j.1469-8137.2011.03771.x, 2011.

Verma, S. B., Dobermann, A., Cassman, K. G., Walters, D. T., Knops, J. M., Arkebauer, T. J., Suyker, A. E., Burba, G. G., Amos, B., Yang, H., Ginting, D., Hubbard, K. G., Gitelson, A. A. and Walter-Shea, E. A.: Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems, Agric. For. Meteorol., 131(1–2), 77–96, doi:10.1016/j.agrformet.2005.05.003, 2005.

Wei, S., Yi, C., Hendrey, G., Eaton, T., Rustic, G., Wang, S., Liu, H., Krakauer, N. Y., Wang, W., Desai, A. R., Montagnani,

L., Tha Paw U, K., Falk, M., Black, A., Bernhofer, C., Grünwald, T., Laurila, T., Cescatti, A., Moors, E., Bracho, R. and Valentini, R.: Data-based perfect-deficit approach to understanding climate extremes and forest carbon assimilation capacity, Environ. Res. Lett., 9(6), 65002, doi:10.1088/1748-9326/9/6/065002, 2014.

Zeller, K. and Hehn, T.: Measurements of upward turbulent ozone fluxes above a subalpine spruce-fir forest, Geophys. Res. Lett., 23(8), 841–844, doi:10.1029/96GL00786, 1996.