## Fluxnet Synthesis Dataset Collaboration Infrastructure

Deb Agarwal (LBNL), Marty Humphrey (UVa), Catharine van Ingen (Microsoft), Norm Beekwilder (UVa), Monte Goode (LBNL), Keith Jackson (LBNL), Matt Rodriguez (LBNL), and Robin Weber(UCB)



The Fluxnet synthesis dataset originally compiled for the La Thuile workshop contained approximately 600 site years. Since the workshop, several additional site years have been added and the dataset now contains over 920 site years from over 240 sites. A data refresh update is expected to increase those numbers in the next few months. The ancillary data describing the sites continues to evolve as well. There are on the order of 120 site contacts and 60 proposals have been approved to use the data. These proposals involve around 120 researchers. The size and complexity of the dataset and collaboration has led to a new

approach to providing access to the data and collaboration support and the support team attended the workshop and worked closely with the attendees and the Fluxnet project office to define the requirements for the support infrastructure.

As a result of this effort, a new web site (<u>http://www.fluxdata.org</u>) has been created to provide access to the Fluxnet synthesis dataset. This new web site is based on a scientific data server which enables browsing of the data on-line, data download, and version tracking. We leverage database and data analysis tools such as OLAP data cubes and web reports to enable browser and Excel pivot table access to the data. The data cubes provide organization and aggregation of data along dimensions such as time to



allow easy retrieval of daily, monthly, and yearly aggregated values. In addition, we have leveraged available collaboration technology (SharePoint) to provide web pages describing the dataset, provide proposer and site PI data access, track data versions and updates, notify users of data updates, and enable support for contact between site PIs and researchers hoping to use their data.

The Fluxdata web site makes available extensive information about the dataset most of which is available under the "Dataset Information" tab. Here you will find interactive site maps, basic site

information, average values, and data availability information as well as descriptions and explanations of the dataset variables and data levels. These reports are presented as spreadsheets for ease of download and are derived from a database containing the Fluxnet data and are refreshed periodically. A list of the approved proposals along with the coordinator and collaborators on the proposal as well as the sites planned for use in the analysis. Links to network specific information are also provided. Currently

	Dataset Ind	armathing Hog (mer Support) Salad	-	Late Browing					190.200			E SHIT
-	Lettrade 2	Dataset Information + Dataset Day 18			restore							-
	Cumu	lativeValuesByVegTy	pe									
a t	acel Web 1	Scons - CanadativeValuesByVeg?	Abe									
Cipier	(*.) (pd	n · Maria										
	A	B C	D	1 B	1	0	el 1		K L		N	
	Cumu	lative Annual Variable	a Va	have but	Cito							
1	Cumu	adve Annoar vanabie	C 10	ues by	Sice							
		and Excel										
4												
4	Working	Group: All Current Fluxnet C	vilac									
ă.												
1				Frack	Rd	RE	Re	SMC	Ta	TR		
				mm y 1	0.0-1.4-1	GI #+2 y-1	G m 2 r 1	m-2 m-2	~	14		
	AT-1984	Austria - Reustry Studies Volley	CRA.	1,284.3229		4.2904	1.3957	71.0731	0.7850	8.4992		
10	AD-How	Australia - Haward Springs	WEA	1,634,2953		6.7934	4.05%	10.3777	32:8205	30.0608		
11	All-Ture	Audrala - Tunberumba	194	1.123.3568		5.5626	2.8324	harris	9.3362	10.4306		
12	AD-Wet	Australia - Wallaby Creek	190	342.5000	0.6604	3.0775	2.1241	23,9139	12:6425	0,4250		
17	80-918	Balgium - Brasschost (Ire Inslag	MF	872.5947		2.5614	1.6303		113,3368	10.5127		
74	RE-78	Reidum - Johney	MF.	1,633.4386		3.2293	1.3362	34,2316	7.8948	9,4191		
15	HE-LON	Belgum - Lonzee	010	453.7293		3.0493	1.0859	28.0118	10.8849	11,9498		
16	BD-Me	Pelgum - Velseim	MF	\$21,8230		3,4672	2.5600		3.3150	(8.2217		
17	BR-Ban	Brazil - Ecstone Benaral Island	TRO	1,002.1673		5.2815	3.1911	44.987)	26.4320	13.0225		
18	BR-Cas	Brazil - Castuana Forest-Almeirm	CBT	1,465.3309		3.5733	4.8806	38.9364	28.0343	25.1231		
15	80.31	Brazil - Rund, - Faz, Nossa Sarborz-	1 GIA			5.8743			24.7717			
20	80-32	Brazil - Rond - Rebio Jeru Ji Ferane-	TRD.	1,646.9595		5.8398			25,2900			
21	BR-Ma2	81828 - Merstus - 252 K34	285	3,038.49(.5		3.0621	14.1181		25-8418			
22	BR-Sel	Brazil - Sankatem-Kmt2-Primary	. CSP	1,592.8210		4.9136	4.0455		25,1605	15.4571		
23	IR-Sal	Brazil - Sankaram-Ken77-Pasture	CHQ	1,284.4250		6.3279		25.0096	28-0437	26.7763		
24	FR-583	Brazil - Santarem-Km83-Looped	184	1,044.9477		3,4836	3.3737	98,2372	25:9964			
25	BR/501	Brazil - Sas Paulo Certado	180	952.6175		6.8634	(4.2365	0+65.55	22.6573			
26	BHP-GNG.	Buttiveria - Chanzi Gress Site	SAV	0.0000		3.1845	0.2689		23.2471	28,8904		
27	BID-CHI	Botavasa - Chargi Maid Sita	WSA.	0.0000		3.1228	0.3176		28-3237	29.7953		
26	BHC-Mat	Butsvere - Neur- Hopere	WSA	258.8420		7.7240	(4.2983	3.2673	22.8255	26.4/27		
25	DICAL	Canada - British Calastola-	EM <sup>2</sup>	1,200-8369	6.3203	2.9407	(2.3811	15.5982	18.6743	0.1097		
30	CR-CK2	Canada - British Calundia-	314	1,089.4367		3.6763	(1.7459	21.3245	29,3447	9,8047		
31	Ck-Call	Canada - British Calundia-	EMP.	1,453-3679	3.5726	3.9050	1.7881	23.1455	9.9056	9.0024		
100	Ch-Gro	Canada - Ortano- Groundhop River-	- 34	432.1527	2,1083	3.6047	1.9385		1,8375	6.1053		
36	Circuit	Canada - Lathbridge	GIA	342.58948		4.4724	2.2197	24.2956	16-8556	7,4672		
11						the second	17 1914 1		THE REAL PROPERTY.	the excelute		
13 14	Ch-Mer	Canada - Eastern Peetland-Mar	O1H	784.6193		0.4822	2.0041					

only the AmeriFlux network has information posted but other networks are welcome to contribute

information as well.



The "Blog" contains news items about the web site, dataset, and processing progress. The "User Support" tab provides access to a user manual describing how to use each of the areas of the site.

Authorized members of the Fluxnet collaboration can sign-in at the web site and obtain access to a range of additional functions. The "Data Browsing" tab provides access to all available site ancillary data as well as summary annual values for the variables. It also provides

direct connection to the data cube for direct browsing of the data. The "Data Download" tab enables download of site years selected by the user. Changes to site ancillary information can also be submitted via the "Ancillary Data Update" tab. Proposal coordinators are able to maintain the list of sites in use for their analysis and communicate with site PIs via e-mail. The next major functionality planned for the Fluxdata web site is to add site PI access to flux-met and ancillary data from their site and support for communication with proposal teams using the site.



Using the Fluxdata web site and Scientific Data Server, researchers can keep up-to-date with the latest



updates to the dataset, select sites of interest for analysis, and communicate with others in the collaboration. The resulting site and capabilities we believe allow individual researchers to concentrate on science rather than data management. This site and server are brought to you by the Berkeley Water Center and University of Virginia in collaboration with the Fluxnet community and in particular Dario Papale. This site is continually evolving and we welcome comments and input from the community. You can reach the support team at <u>fluxdata-support@fluxdata.org</u>.

The research and development described in this paper is funded by Microsoft Research's eScience program. This work is also supported by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

This article appeared in the FLUXNET newsletter: FluxLetter, Vol 1, Issue 1, 2008, pp. 5-7.