Combined AmeriFlux / NEON breakout session
Joint NACP-AmeriFlux PI Meeting 2017

Wednesday, March 29, 2017, 10:00 - 11:30am EDT

Due to overlap in topics and participants, the breakout chairs decided to merge the two sessions:

- AmeriFlux for the broader carbon cycle research community: New data products and future prospects (Chairs: Trevor Keenan, Margaret Torn, Deb Agarwal, Gilberto Pastorello, Room: White Oak A)
- Catalyzing carbon cycle science through synergies among research networks (Chairs: Stefan Metzger, Dario Papale, David Durden, Ankur Desai; Room: White Oak B)

Table of Contents
1  Questions asked..................................................................................................................2
2  Graphical summary ..............................................................................................................2
3  Detailed notes ......................................................................................................................2
   3.1 Opening remarks...........................................................................................................2
   3.2 What current data products and capabilities are most beneficial to your work using carbon flux data and are you interested in helping to develop them? ..................................................3
   3.3 What products and services are needed to support data collection and data contribution? ..........................................................................................................................3
   3.4 What are the most significant missing products? ...........................................................3
   3.5 Are there other data, services, and algorithms that are available/needed? ....................4
   3.6 What resources for / approaches to synergy should networks be focused on to gain the greatest benefit to the science community? ................................................................5
   3.7 What priority should be given to the resultant outcome? ..............................................5
   3.8 How best to encourage community development and input for the development of these tools and resources? ......................................................................................5
4  Deliverables.......................................................................................................................5
1 Questions asked

1. What current data products and capabilities are most beneficial to your work using carbon flux data and are you interested in helping to develop them?
2. What products and services are needed to support data collection and data contribution?
3. What are the most significant missing products?
4. Are there other data, services, and algorithms that are available/needed?
5. What resources for / approaches to synergy should networks be focused on to gain the greatest benefit to the science community?
6. What priority should be given to the resultant outcome?
7. How best to encourage community development and input for the development of these tools and resources?

2 Graphical summary

The participants provided detailed input (see Sect. 3), which generally encompassed the 6 broad areas above.

3 Detailed notes

3.1 Opening remarks

- AmeriFlux
  - Overview of AmeriFlux / FLUXNET data products and plan for future releases
  - Ameriflux sites: CDIAC site going way, BASE products replaces L2, working on BADM, all existing, processed BADM data in BIF files,
  - Fluxnet2015 - open-source code in development (focus on GPP/RE partitioning algorithms) by early 2018, then one last Fluxnet set, then 5 yrs
- NEON
  - overview of NEON site info and data products (how to see what’s available) and timelines
- Field site link http://www.neonscience.org/science-design/field-sites tells you about construction status and link to data
- Data portal http://data.neonscience.org/home is object based, contextual themes, lots of products on soil/met available now
- L1 state variables out in June, flux variables end of 2017
- Protocols/instrumentation down the road. Community input sought/ workspace
  - Ingest NEON data into AmeriFlux, pilot site and hopefully all others; moving collaboration forward

3.2 What current data products and capabilities are most beneficial to your work using carbon flux data and are you interested in helping to develop them?
- Trevor: Essential metadata lat/lon easily accessible when download alongside flux data
- Luke: Soil respiration data alongside flux; e.g. NEON collection soil CO2 gradient, need to algorithm development

3.3 What products and services are needed to support data collection and data contribution?
- Data structures: how to accommodate data that do not easily fall into current format (different time resolution, spatial data)
- How to incorporate sites that have rich data (e.g. profiles, soil, hydrology), but not EC measurement per-se
- AmeriFlux - are they open to new sites and new types of cover (coastal, aquatic, coastal blue carbon, agricultural sites, sites with non-CO2 GHG fluxes, non-Americas data (Africa) and other short-term flux data in Americas, make it North America + South/Central America, beyond US.
- Gilberto: Automated method of flux classification for metadata
- Chris G: standardized suite of BADM products that are prioritized
- Inke: Flux towers at LTER -> many are not in AmeriFlux. Lots of coastal sites coming online -> what “class” to put them -> wetland or water?
- Tagir: No flux towers in Nevada, then a paper shows up, how to get AmeriFlux (Deb: we are still a coalition of willing. International: Europeans handle non Americas)
- Mallory B: footprint analyses in papers, ways to upload
- Andy S: WE have Excel files of data. Interested in depositing, but not have time to convert to standardized format. Flexibility to add other data sets
- Sebastien: needs to be 2-way interaction, we hear you, we’re doing the same work (we also cringe when we see the BADM, but we need to have documentation and standardization); this is the coalition of the willing (we can assist, but not force)
- Very few criteria for joining AmeriFlux. If you have data, send it. Provide enough information on data
- Deb O’Dell: FLUXNET, has data in Africa, short term data in Ohio, short term data in Tennessee; imaging important especially in agriculture (LAI or other parameters).
- BADM data: add flexibility to input data; NEON coming from ecological side can provide solution for data reporting

3.4 What are the most significant missing products?
- Standardization of terminologies, incl. both flux and metadata; mitigate website confusion
- Consistent processing (incl. QAQC) from raw data to 30 min
- Hope for NEON to supplement ecological / rich contextual data
- Footprint remote sensing imagery; spatial representation of data; plot it on map, where soil chambers, etc. within a site
- Snow cover information
- Higher time-resolution of remote sensing products
- Accessing / filtering sites based on instruments
- Missing data. Knowing % of data available by year
- NEON AOP: thread data server, pixel-by-pixel uncertainty
- Flux towers are not evenly distributed, prioritize where to get data to fill sampling gaps
- Paul H: Belowground availability
- High-freq data available. Multiple heights available. Download or PI?
- Uncertainty in partitioning - defining low-light period for respiration in high-latitude
- Search features
  - general
    - Web of science like search, search by index terms, semantic search
    - Portal: querying by actual data availability, could be as easy as showing a variable-over-time plot
    - Dave Johnson: if there’s a way to filter sites based on instrumentation used, to allow notices to be sent to community about instrument fixes
  - AmeriFlux
    - Site search exists in AmeriFlux. Site general info is needed to make it work.
    - Site sets capability exists. Can now save a search as a “set”
    - Need site PIs to provide and be consistent
    - Tell AmeriFlux what you can’t find. No filter option in download page.
    - Non-CO2 species/data (COS/SIF/13C) -> N2O is there. Just ask, and AmeriFlux can add
  - NEON
    - NEON has created usability tools - R packages for gathering/filtering data from command line (nneo), met scanner R package - given NEON site, find any met network in a radius from that site (ASOS, etc.. incl. AmeriFlux),
    - HDF5 as pathway to attach metadata to timeseries etc.
    - NEON saves search

3.5 Are there other data, services, and algorithms that are available/needed?
- Support with calibration (as a service)
- Calibration requirements on concentration data / use together with RS data?
- Search by site data flux availability (CH4 flux)
- Alexey: NEON AOP, THREDDS server - subsetting remote sensing data, derived products-> pixel level uncertainty, calibration
- Gordon: Data filtering for availability, Metadata that can be updated by users
- Ben R: OCO data entry portal like MODIS cutouts. Tech blogs on solar power, and other instruments. Consistent way to Mean annual precip/temp
- Yun Liu: data assimilation, uncertainty estimates!
- Marcy: 2-way street for communication, box if you see problem with data to let PI know!, cite DOI, data providers love to collaborate
- Make it easier to combine aircraft and tower data
- Open-development
NEON code is gradually being switched public once thoroughly tested, already possible to grant access on request; End of this year NEON tool set, combined R packages end-to-end workflow. Can leverage to improve code/add features. Cross-network repository

- GitHub repo for AmeriFlux code
- Whole intent is open-source and let individuals run code themselves, share, develop
- High-freq: EddyPro + Eddy4R, REDdyProc (for post-processing)
- Repositories exists, need to make public, email Ameriflux or NEON to get involved for now in private repos
- What is method to add to codebase vs making a branch
- Need feedback on tools that help (good, bad, and ugly)
- Code repository for sharing code snippets is high interest. Need to work out licensing.

### 3.6 What resources for / approaches to synergy should networks be focused on to gain the greatest benefit to the science community?

- Integration with networks of other agencies (NEON - AmeriFlux just as a starting point); LTER, ISCN, NOAA, CZO, general philosophy for sharing, protocols for data (like registering formats with WMO)
- Integrated data access point, making data more available across networks
- Improved data availability browsing capabilities
- Web tools for data crunching and gap filling, high-freq and half-hourly
- Tash: Phenology, phenocam, FaPar, NDVI, soil moisture, all useful, better ways to facilitate collaboration
- Peter: Rants. Contracted sites are one thing. But AmeriFlux is broader - coalition of willing. NACP is both. PMs tell PIs if they are in NACP, told you must archive data. NACP can help get data from funded programs/ agencies. Can help investigators submit data.
- Stefan: Leading flux team at NEON. Common definition of NEE (turbulence + storage term at all sites?). NEON planning to provide spatial representativeness / location bias will be provided by NEON. Common strategy for how that could be useful

### 3.7 What priority should be given to the resultant outcome?

- Dennis: emphasize 2-way communication!; but also careful with standardization (Ray Leuning’s “know thy site” is more important), look to long tail of your data; documentation of the site is important (easy way to do that are important)

### 3.8 How best to encourage community development and input for the development of these tools and resources?

- Tess: short-list open source packages/tools for working with EC/AmeriFlux data and who is maintaining them
- Rosie: used to FLUXNET data; if any sites have phenology (e.g., phenocam) and soil moisture data; as a modeler don’t want to just take people’s data but collaborate (what tools/suggestions on how to support collaboration)

### 4 Deliverables

AmeriFlux and NEON will use this information to continue working towards:

1. List of desired carbon flux products and capabilities.
2. List of volunteers to contribute algorithms.
3. Prioritized list of resources, methods, and tools that are needed to synergistically address emerging questions in carbon cycle science.
4. Owners to work toward implementation of action items.
5. Collaborative workspace to share ideas and track progress on identified items.