Stefan Metzger[1], David Durden[1], Andy Fox[2,3], Greg Holling[1], Hongyan Luo[1], Natchaya Pingintha-Durden[1], Cove Sturtevant[1]

[1]: National Ecological Observatory Network, Fundamental Instrument Unit, Boulder, Colorado, USA

[2]: Arizona State University, Tucson, Arizona, USA

[3]: National Center for Atmospheric Research (NCAR) , Boulder, Colorado, USA

Acknowledgements: NEON Technical Working Groups; Ankur Desai, Tammy Halstead, Mark Kessel, Natascha Kljun, Andrei Serafimovich, Matthias Sühring, David Weinstein, Ke Xu, Liangying Zhang

NEON's first release of eddy-covariance data products and software tools

National Ecological Observatory Network

A project sponsored by the National Science Foundation and operated under cooperative agreement by Battelle.

the spirit

- incorporate lessons-learned through collaborations with bottom-up networks like AmeriFlux, ICOS, LTER, TERN…

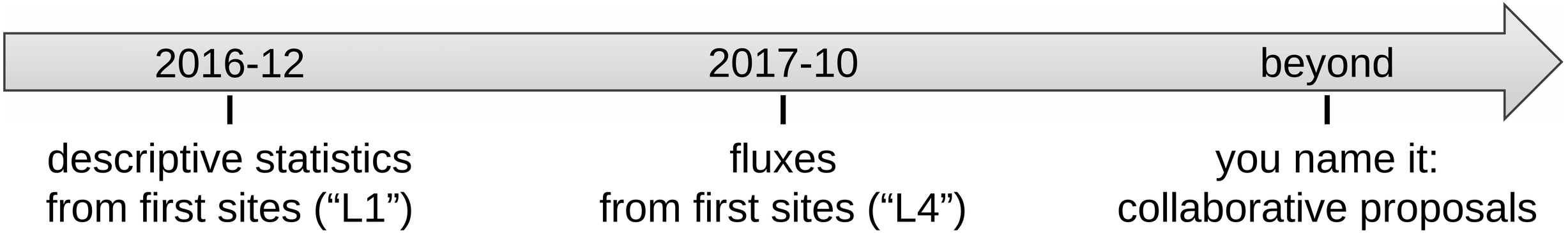
- NEON's centralized approach lends itself to explore novel systemic solutions

- starting to give back:
  - eddy-covariance data products
  - eddy-covariance R-packages
  - eddy-covariance usability tools

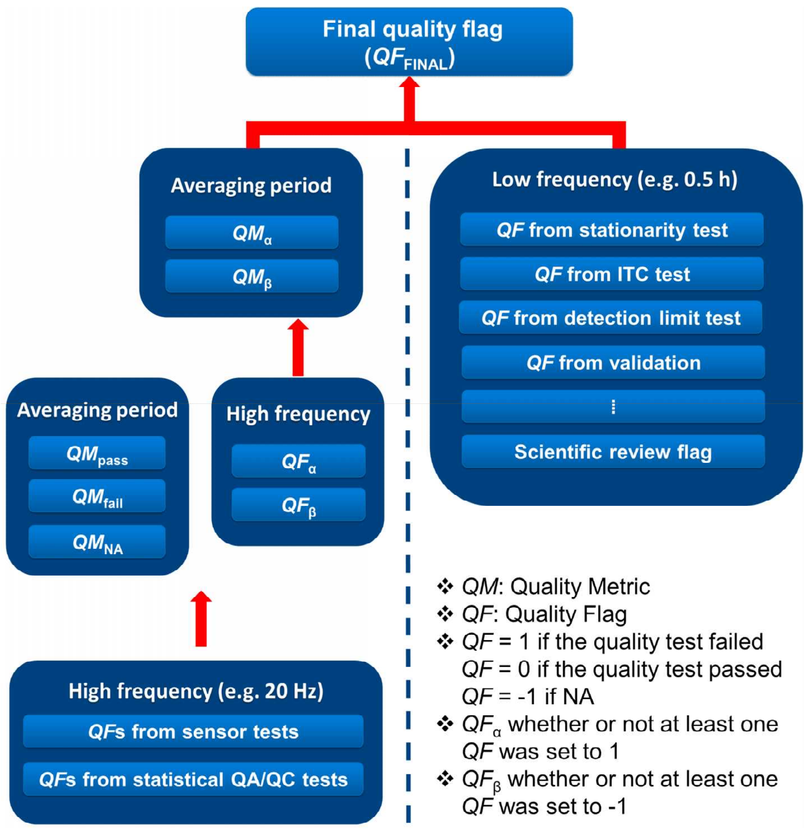
credit: http://www.christianpost.com/news/give-and-take-37054/

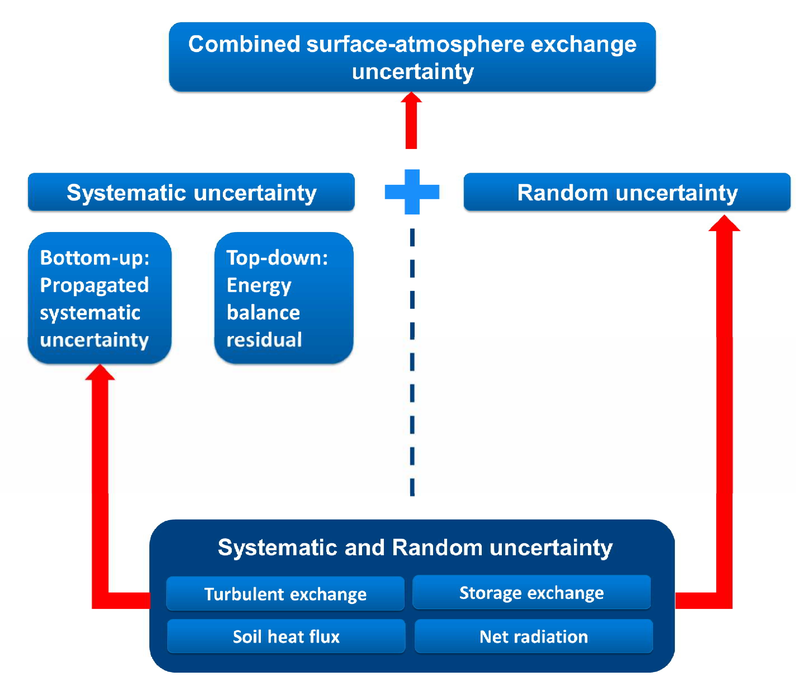
- synergize ongoing research efforts across science communities: **2:40 pm breakout** "Processing best practices and methods including tools and workflows"

eddy-covariance data products: contents and schedule



- selectable formats: basic (order 100 variables), expanded (order 1000s variables)
- including data quality and uncertainty budgets (poster: Pingintha-Durden et al.)





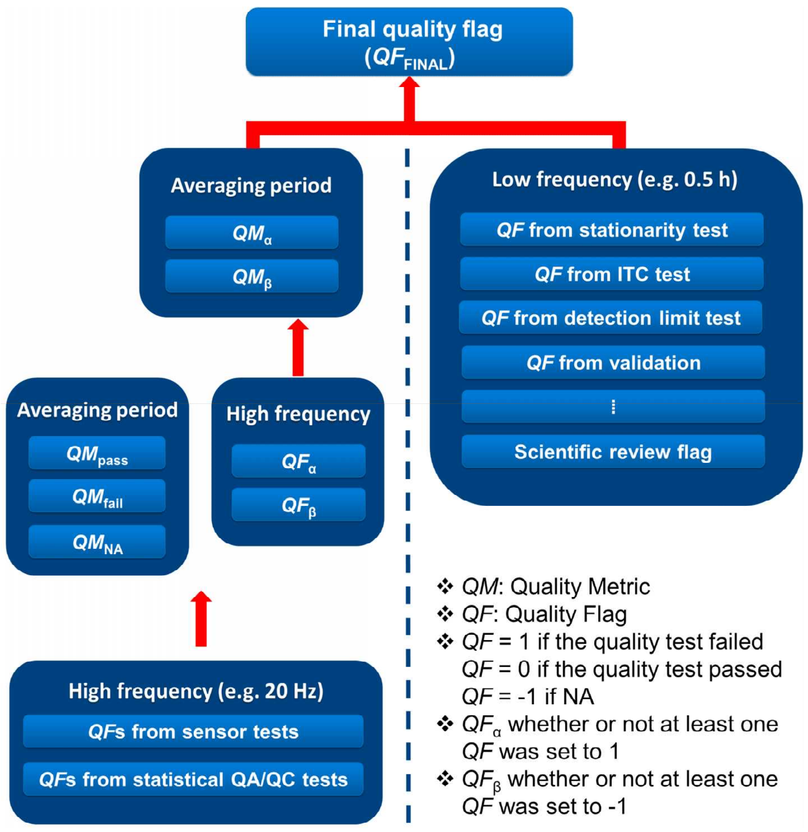
Battelle
The Business of Innovation

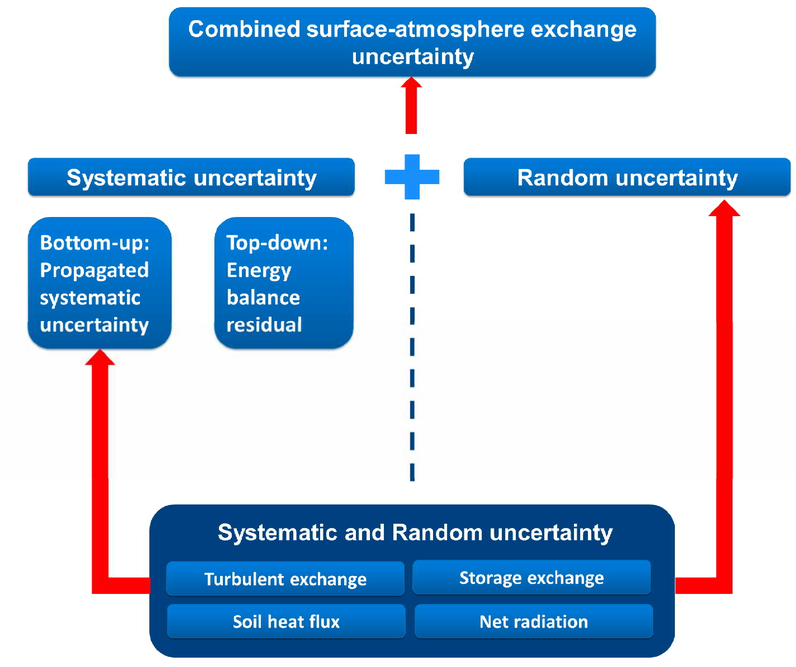
eddy-covariance data products: contents and schedule

2016-12	2017-10
descriptive statistics from first sites (“L1”)	fluxes from first sites (“L4”)

PROGRAM SOLICITATION
NSF 16-521
MacroSystems Biology and Early NEON Science
collaborative proposals

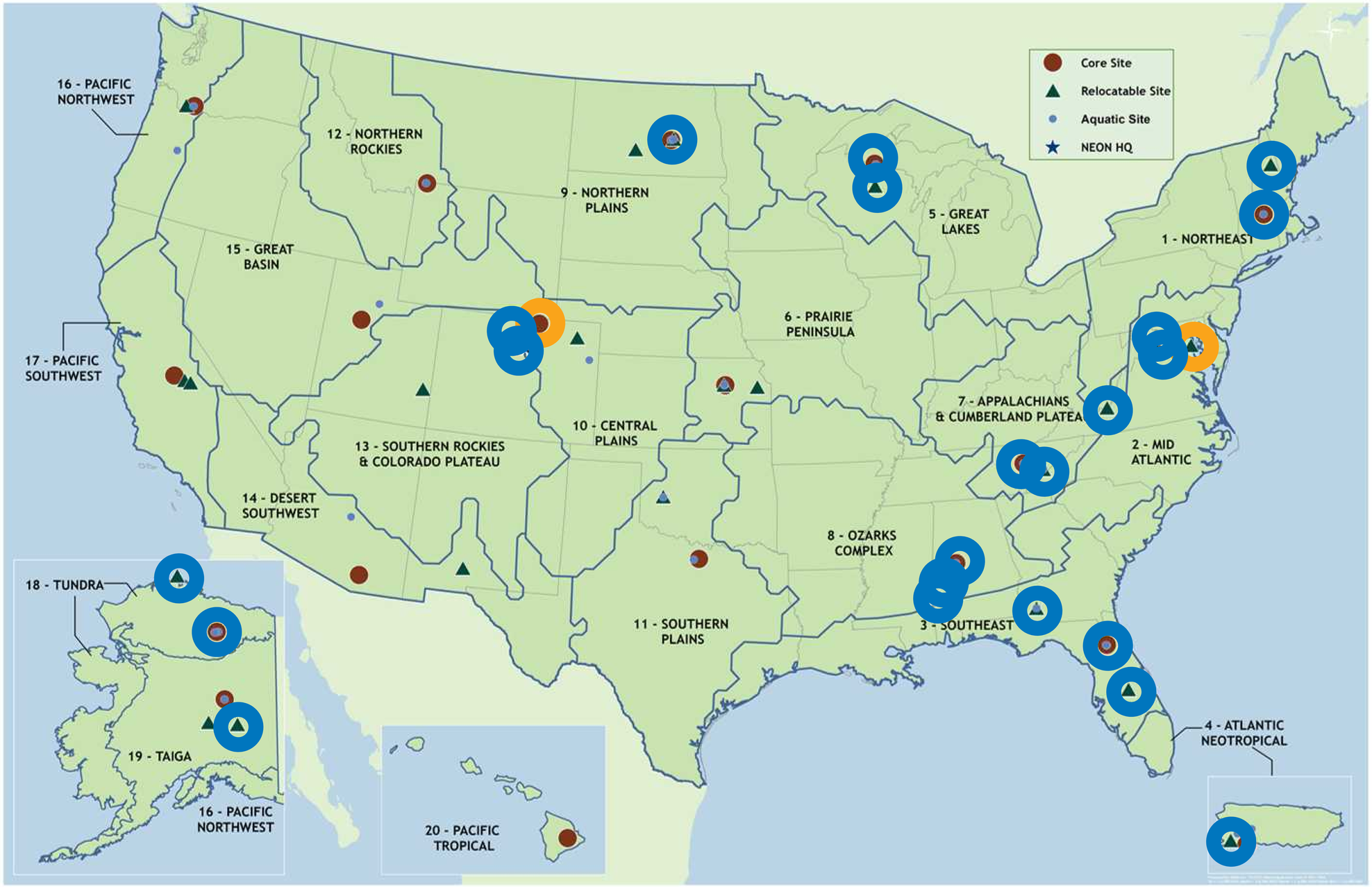
- selectable formats: basic (order 100 variables), expanded (order 1000s variables)
- including data quality and uncertainty budgets (poster: Pingintha-Durden et al.)





Battelle
The Business of Innovation

eddy-covariance data products: sites and schedule



- +0 months: 2 sites
- +6 months: 25 sites
- +12 months: all 47 sites
- provisional data until first versioning (mid-2019)

eddy-covariance data products: how to access

- neonscience.org/data-resources
  - calibrated raw data (“L0p”): per request (web-browser)
  - data products (“L1” … “L4”): per data portal (web-browser and API)

- file format: Hierarchical Data Format (HDF5, poster: Durden et al.)
  - self-describing incl. all contextual information (“metadata”)
  - discoverable with standard graphical and command line software

003_001 at /SERC/L0_irgaPresValiLine/ [HDF5TIS_L0_prototype.h5 in C:\Users\ddurden\Sh

Table

	SAMPLE	DATE	TIME	Delta.T.s.	Julian.Day.l...	PRESSUR...
0	1	2016-04-22	00:00:01.1...	1.1226267...	112.75001...	10.69681
1	2	2016-04-22	00:00:02.1...	2.1226401...	112.75002...	10.68373
2	3	2016-04-22	00:00:03.1...	3.1226532...	112.75003...	10.70403
3	4	2016-04-22	00:00:04.1...	4.1226663...	112.75004...	10.73999
4	5	2016-04-22	00:00:05.1...	5.1226787...	112.75006...	10.71329
5	6	2016-04-22	00:00:06.1...	6.1226916...	112.75006...	10.72485
6	7	2016-04-22	00:00:07.1...	7.1227045...	112.75008...	10.7305
7	8	2016-04-22	00:00:08.1...	8.1227169...	112.75009...	10.72704
8	9	2016-04-22	00:00:09.1...	9.1217298...	112.75010...	10.73781
9	10	2016-04-22	00:00:10.1...	10.121741...	112.75011...	10.73665
10	11	2016-04-22	00:00:11.1...	11.121752...	112.75012...	10.72276
11	12	2016-04-22	00:00:12.1...	12.121765...	112.75013...	10.7347
12	13	2016-04-22	00:00:13.1...	13.121775...	112.75015...	10.75291
13	14	2016-04-22	00:00:14.1...	14.121787...	112.75016...	10.7393

eddy-covariance R-packages: eddy4R

- eddy4R family of R-packages
  - eddy4r.base, eddy4r.turb, eddy4r.qaqc…
  - installable from public Github repository (devtools::install_github())
- focus: from raw data to 30-min
- release schedule follows eddy-covariance data products (2016-12, 2017-10)

Eddy-covariance calculation for R: Base package

Documentation for package ‘eddy4R.base’ version 0.0.18

- DESCRIPTION file.

Help Pages

Conv	Conversion Factors
def.agr.vari.seSq	Determining mean, external, internal and total variance, and squared standard error
def.aply.conv.poly	Apply polynomial conversion
def.bin	Binning data
def.cmpr.out.refe	Compare output against reference
def.coef.corl	Coriolis coefficient
def.conv.az.cart	Decomposing azimuth angles to cartesian vectors
def.conv.body.met	Coordinate transformation from CSAT3 body coordinate system to meteorological coordinate system
def.conv.cart.az	Composing azimuth angle from cartesian vector data
def.conv.unit	Unit conversion

Eddy-covariance calculation for R: Turbulent flux

Documentation for package ‘eddy4R.turb’ version 0.0.10

- DESCRIPTION file.

Help Pages

const_f	63% frequency constant after Aubinet (2012) Eq. 4.22
const_t	63% time constant after Aubinet (2012) Eq. 4.22
COSP.fwd	Generate cospectra
COSP.plot	Plot cospectra
def.dist.rgh	Aerodynamic roughness length
def.func.univ	Integral over the universal function
def.itc	Integral turbulence characteristics
def.nois	Determination of noise and detection limit for eddy-covariance turbulent fluxes
def.stna	Stationarity tests
find_F0	Determine cutoff frequency empirically
find_FX_og	Determine spectral peak using an Ogive method

Eddy-covariance calculation for R: Quality assurance and quality control

Documentation for package ‘eddy4R.qaqc’ version 0.0.7

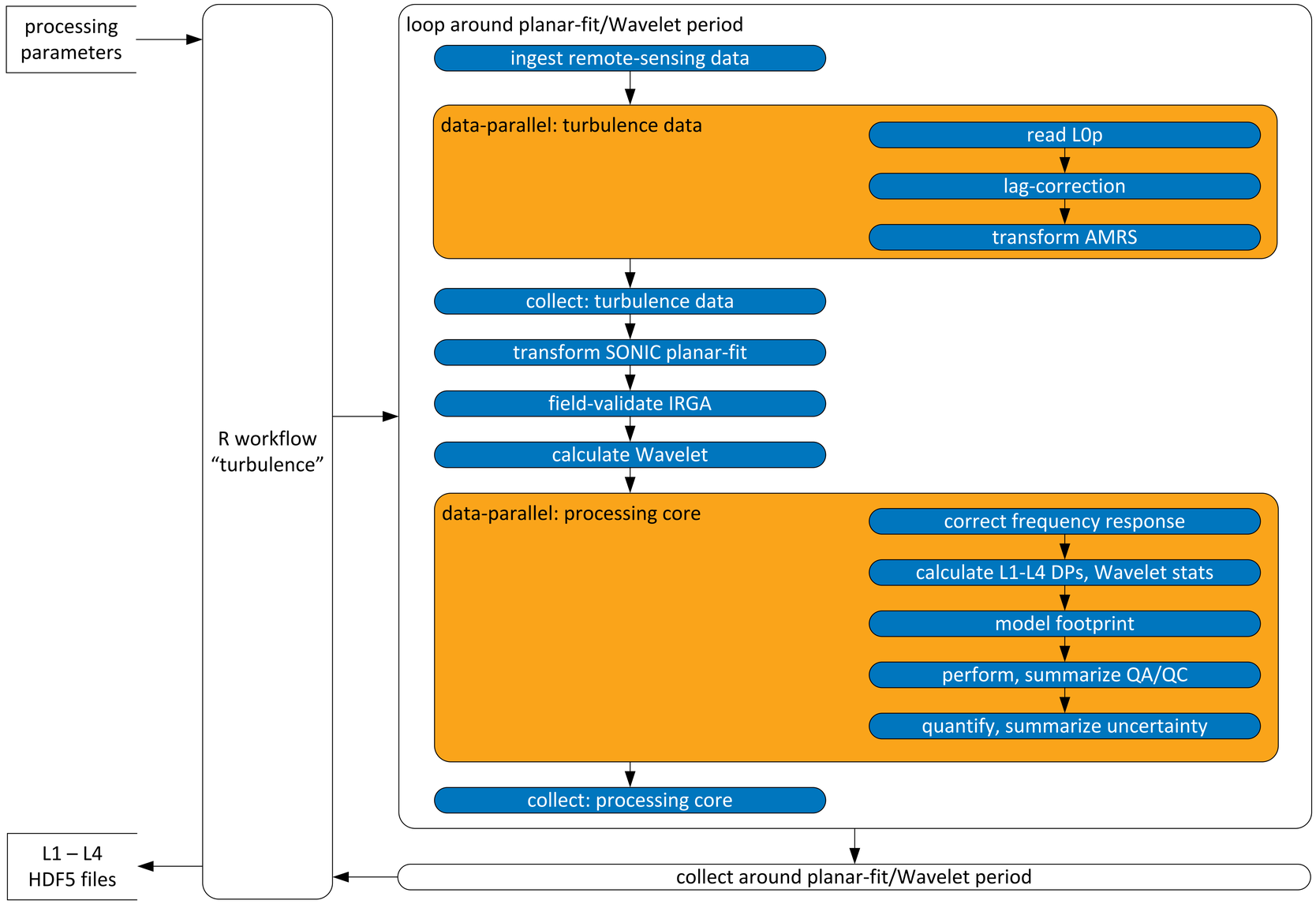
- DESCRIPTION file.

Help Pages

def.dspk.wndw	Determine spike locations using window-based statistics
def.mapp.fulc.form.key	Map data names to internal keys in Fulcrum form schema
def.plau	Plausibility tests (Range, Step, Persistence, Null, Gap)
def.plot.qfqm.l1	Plot quality flags and quality metrics (basic L1 data products)
def.qf.finl	Final Quality Flag (basic L1 data products)
def.qfqm.l1	Quality flags and quality metrics (basic L1 data products)
def.qm	Quality Metrics

eddy-covariance R-packages: own scientific workflow

- modularly adjustable and extensible workflow templates, calling…
  - …wrapper functions
    - …definition functions



eddy-covariance R-packages: advances for sci. discovery

- designing end-to-end scientific analyses in single, accessible interpreter language
  - NEON's eddy4R: open-source modules for raw data processing
  - MPI's REddyProc: open-source modules for 30 min aggregate processing

- eddy4R unique features
  - computationally efficient: adaptive single-pass workflow, fully parallelized
  - calculation of storage flux from tower profiles (poster: Luo et al.)
  - alignment and motion compensation for wind measurements
  - processing data from towers and moving platforms (aircraft, buoys etc.)
  - time-frequency decomposed fluxes (e.g., Vaughan et al, 2016)
  - multi-dimensional environmental response functions (e.g., Metzger et al., 2013)
  - regionalization at high spatio-temporal resolution (e.g., Xu et al., 2017)

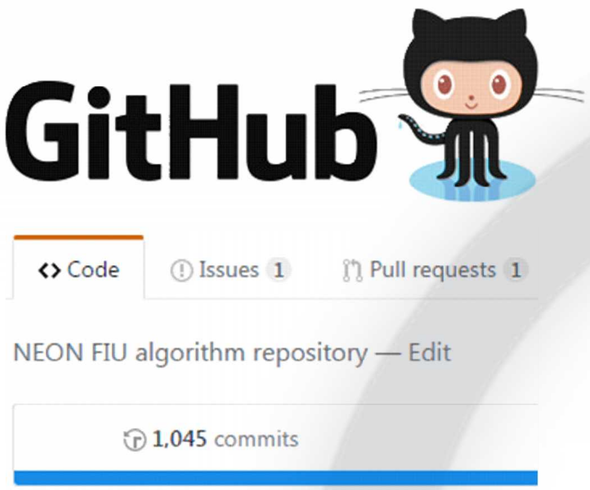
eddy-covariance usability tools: eddy4R-Docker image

- Docker = shipping container system for code (poster: Holling et al.)

- “containers wrap a piece of software in a complete filesystem that contains everything needed to run: code, runtime, system tools, system libraries”
  - efficient: shares host operating system (OS) instead of guest OS emulation
  - reproducible: same results, regardless of the host operating system
  - lightweight, distributed via a web-based portal (hub.docker.com)
  - deployable at scale, from laptop to massively parallel applications

eddy-covariance usability tools: DevOps cycle

(3)



(4)

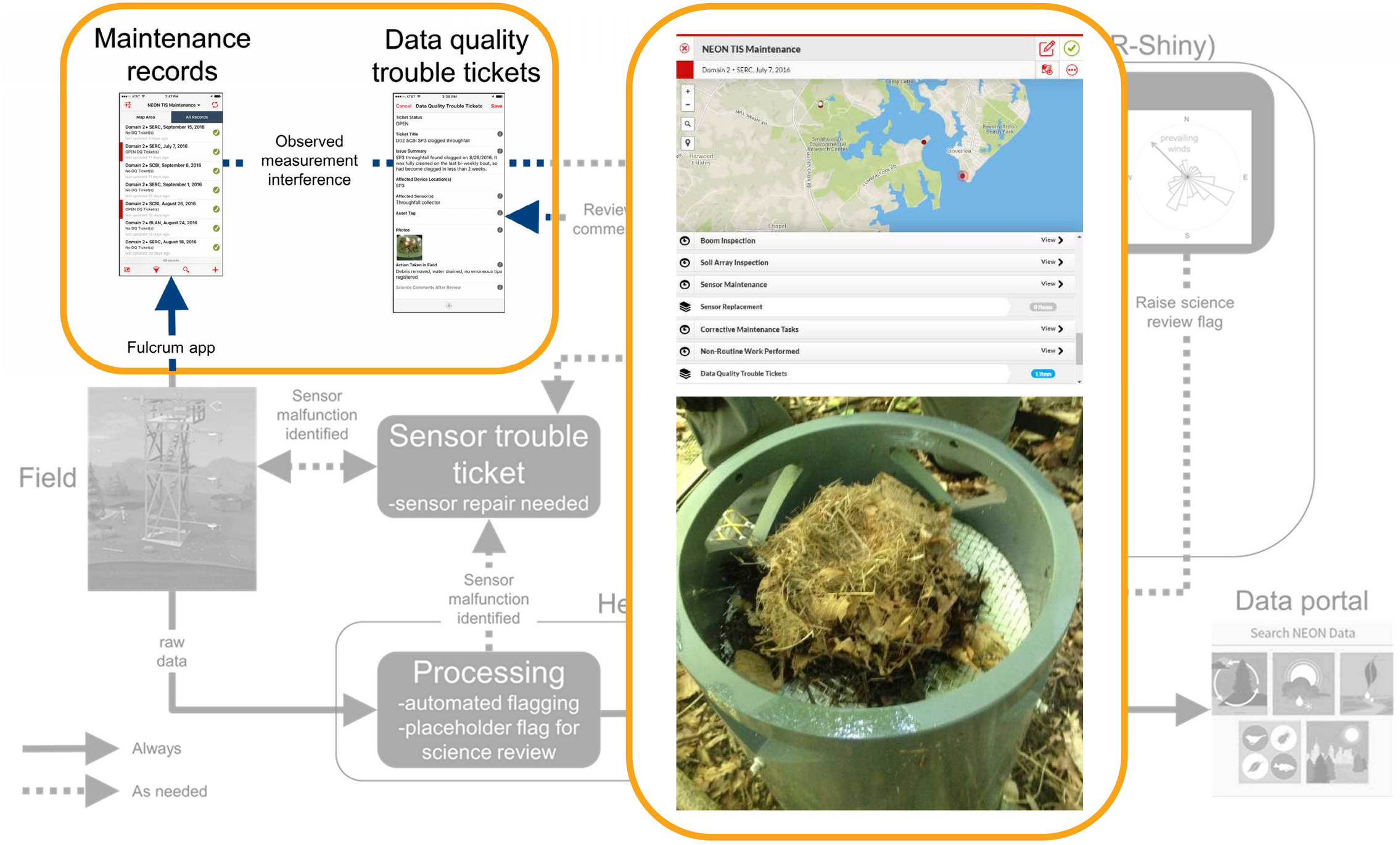
(2) Science

(5) CI

(1) Science Community

eddy-covariance usability tools: operations management

- problem tracking and resolution along the entire chain (poster: Sturtevant et al.)



conclusions

- 2017: first NEON eddy-covariance data, R-packages and usability tools become available

- NEON’s eddy4R + MPI’s REddyProc R-packages: end-to-end scientific workflows in single R-environment

- eddy4R-Docker: turn-key, reproducible, extensible and portable data processing + analysis environment

- Fulcrum: efficient field operations management along the entire data chain

- direct participation in open development: accelerating and leveraging research across communities

outlook

- 2:40 pm breakout "Processing best practices and methods including tools and workflows“
    - synergize ongoing research efforts across science communities

- 4:30 pm NEON posters: Durden et al., Holling et al., Luo et al., Pingintha-Durden et al., Sturtevant et al.

- 2 PostDoc openings
    - storage flux
    - flux uncertainty budget

- coordinated efforts on higher-level processing, value-added data products
    - NEON ↔ AmeriFlux / FLUXNET