

Accessing and Using Sensor Data within the Kepler Scientific Workflow System

Derik Barseghian¹, Ilkay Altintas², Matthew B. Jones¹

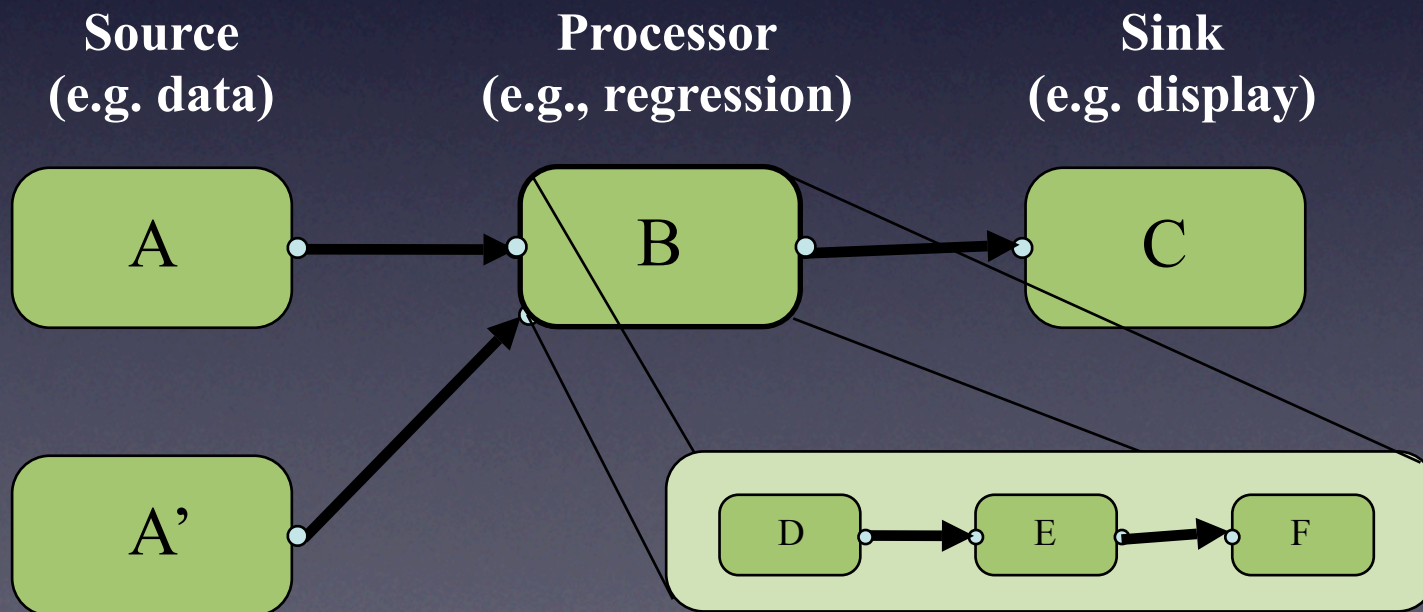
¹ National Center for Ecological Analysis and Synthesis, University of California Santa Barbara

² San Diego Supercomputer Center, University of California San Diego
{barseghian, jones}@nceas.ucsb.edu, altintas@sdsc.edu

Overview

- Scientific Workflows
- Kepler Scientific Workflow Software
- REAP (Realtime Environment for Analytical Processing)
 - Terrestrial Ecology Usecase
 - MET Station
 - REAP Terrestrial Ecology Workflows
 - Example workflow
- Summary and Future Work

Scientific Workflows



Kepler and Kepler Workflows

file:/Users/derik/dev2/kepler_svn/workflows/neon/RBNBToKeplerDisplay4SDF_eim.xml

File Edit View Workflow Tools Window Help

Components Data

Search
composite
☐ Search repository
Sear... Reset

Search Results
Components
General Purpose
CompositeActor
DistributedCompositeActor
Workflow
CompositeActor
DistributedCompositeActor
Run Composite Actor

5 results found.

DDF Director

requestStart: "2008-01-01 00:00:00"
days: 30

Baskett Slough MET Station Datastreams

Plot Air Temperature

Plot Battery Voltage Minimums

execution finished.

REAP

Realtime Environment for Analytical Processing

- Adding infrastructure to the scientific workflow software Kepler to access, monitor, analyze and present information from field-deployed sensor networks.

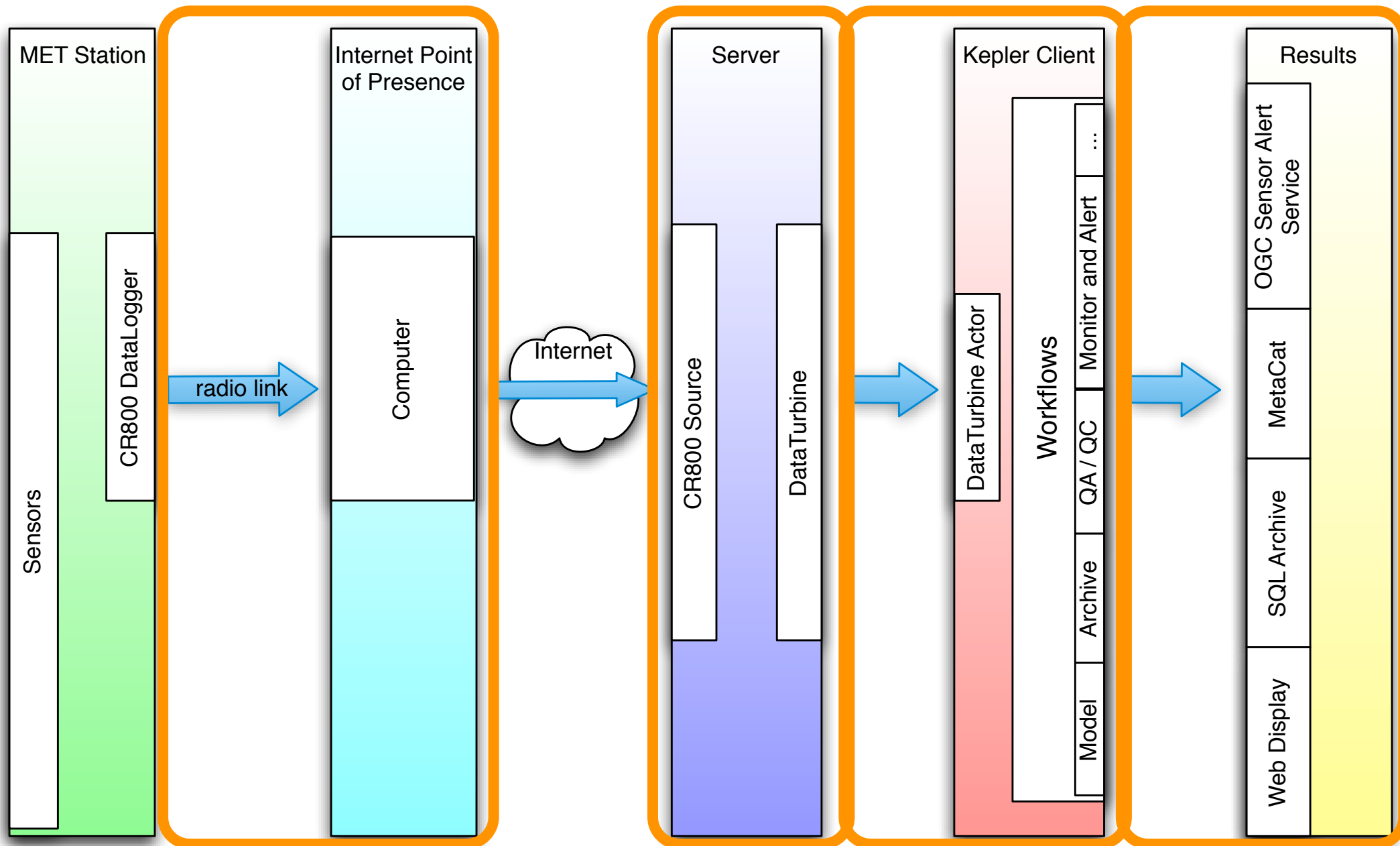
REAP Terrestrial Usecase

Explore the impacts of abiotic factors (light, temperature, rainfall) on the dynamics of plant host populations and their susceptibility to an insect vectored viral pathogen.

This study requires:

- A detailed understanding of grass community phenology -- measure:
 - ambient meteorological conditions
 - soil moisture
 - biomass accumulation in the grass canopy
- Support for the analysis and modeling of sensor data to detect:
 - local thresholds -- e.g. hours exceeding developmental thresholds for aphids
 - long-term trends -- e.g. within- and among-season soil moisture trends
 - significant events -- e.g. timing of peak plant biomass
- Integration of sensor data with archived data to assess the relative impacts of disease, plant composition, rainfall, temperature, and soil nutrients on competitive interactions among grasses

REAP Terrestrial Usecase Architecture



Terrestrial Ecology Usecase Workflows

Three areas of workflow focus:

- 1) event detectors -- alerting on monitored data streams.
- 2) QA/QC-- producing higher-level derived data products to be archived for use in post-hoc analyses.
- 3) post-hoc analyses of data

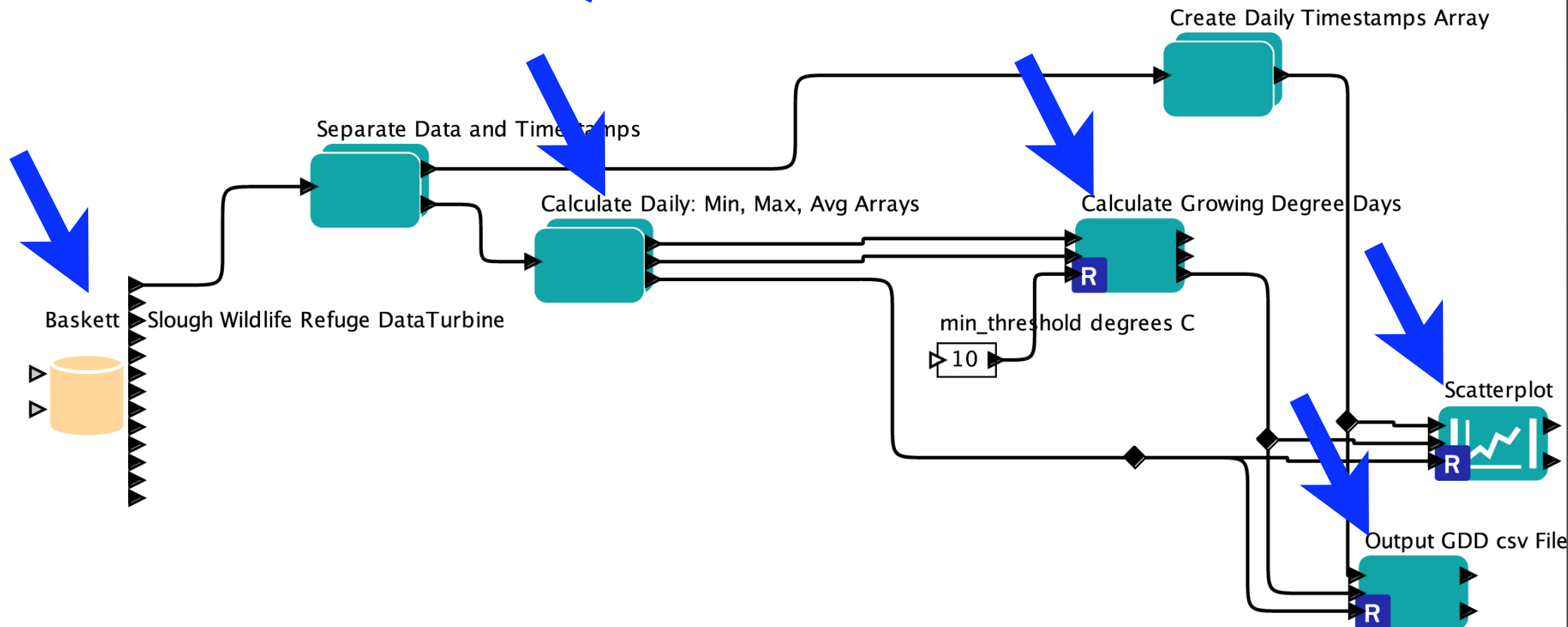
Growing Degree Days from Temperature Data

SDF Director



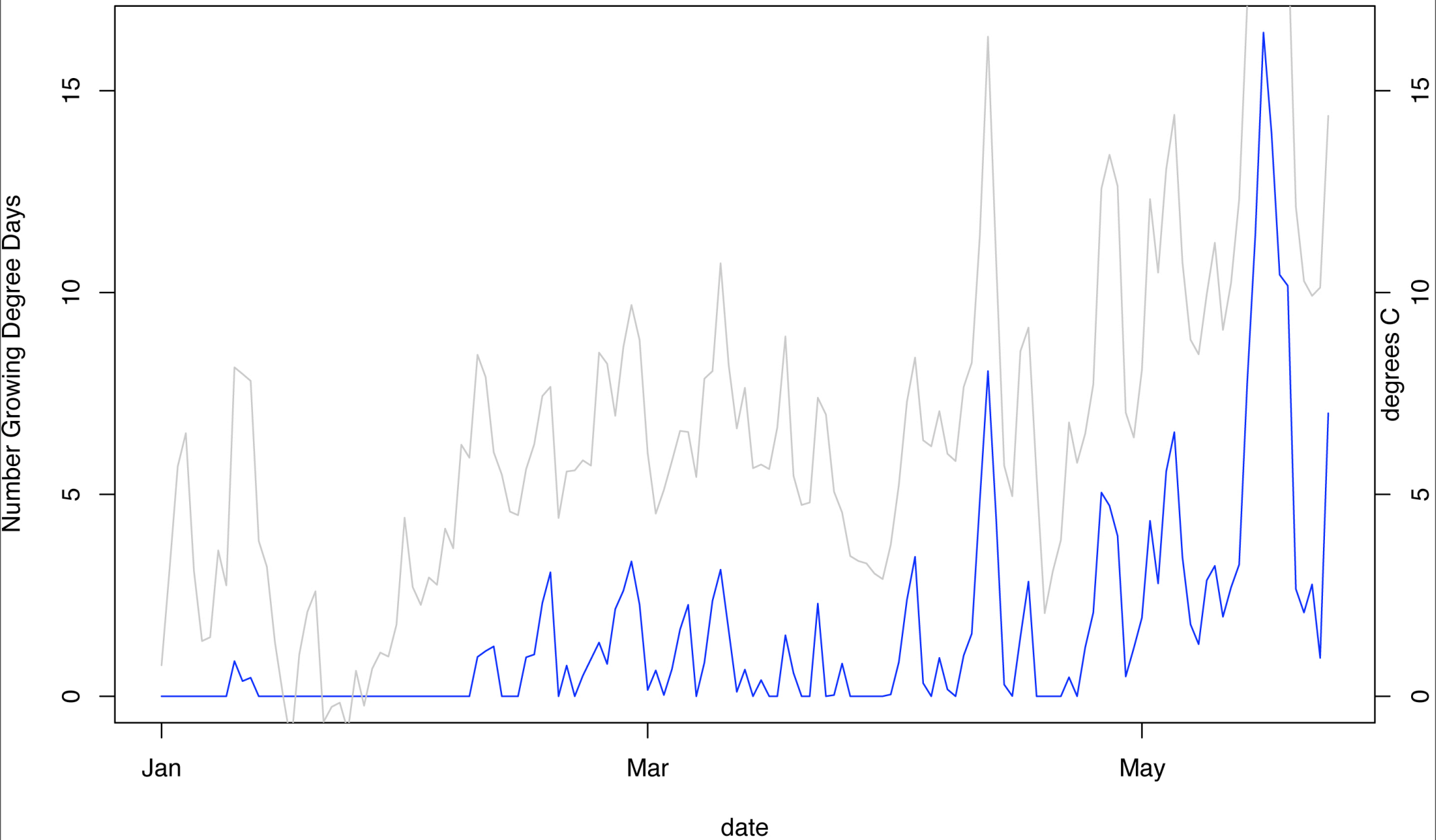
Calculate and Plot Growing Degree Days using temperature data from REAP01 station at Baskett Slough.

- requestStart: "2008-01-01 00:00:00"
- days: 145
- requestDuration: days*24*60*60
- sampleInterval: 600
- maxLengthOfDataArrays: requestDuration/sampleInterval

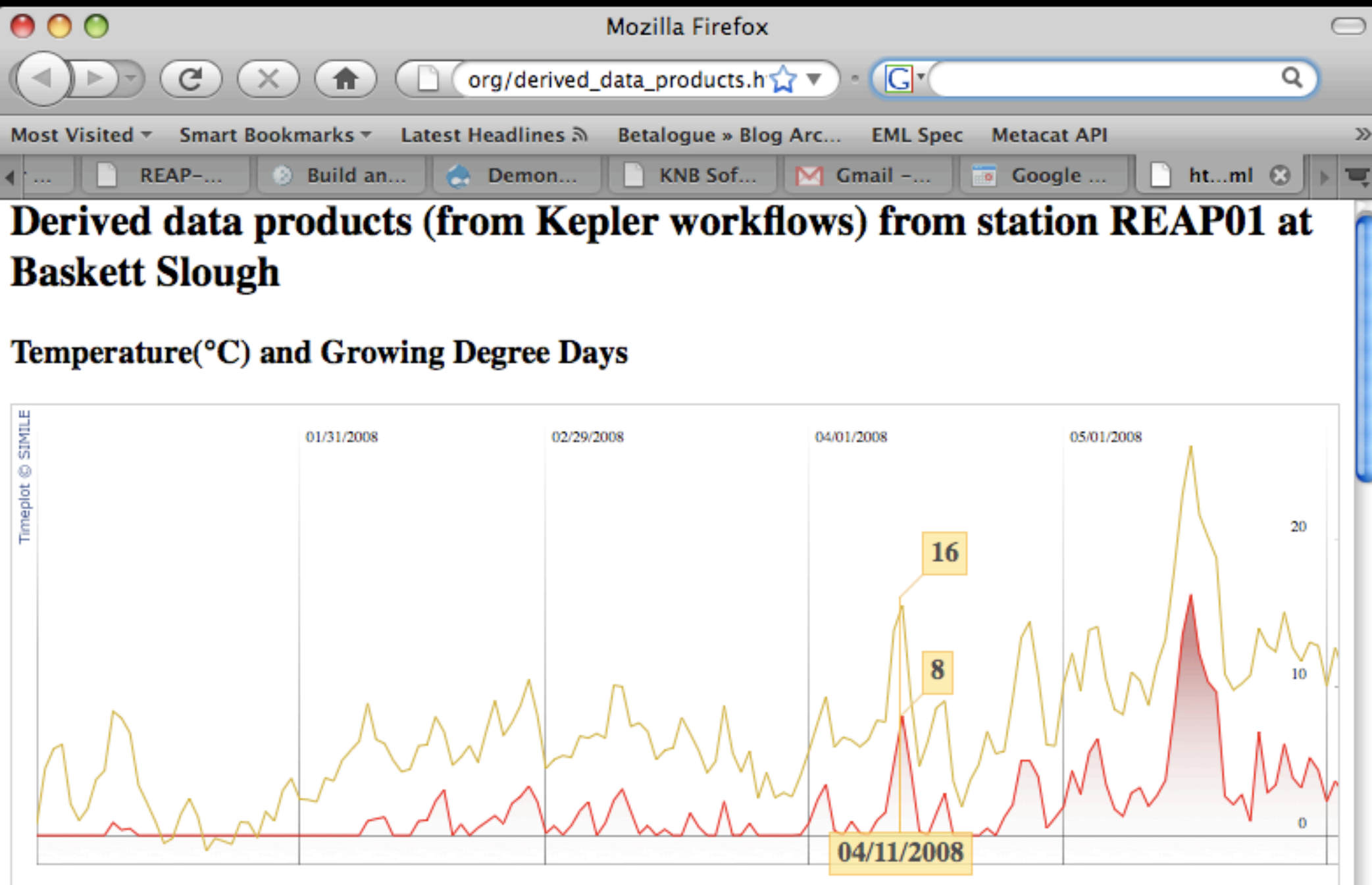


Kepler Plot

Temperature and Growing Degree Days at REAP01



Online plot



Summary and Future Work

REAP is adding infrastructure to Kepler to access, monitor, analyze and present information from field-deployed sensor networks:

- Access data from technologies used by environmental observatories, like DataTurbine, OPeNDAP* (*see oceanography poster), Antelope.
- Interfaces and displays to browse data and be alerted when events of interest occur.
- Interfaces to format, filter and archive datastreams, into data-stores like MetaCat and SQL Databases, using relevant standards like KNB's EML, OGC's Sensor Web Enablement TransducerML
- Provide a coherent environment within which to use sensor data in conjunction with existing datasets
- Work towards bidirectional communication (e.g. sensor control, e.g. sampling rate changes from within Kepler)

Acknowledgements

This material is based upon work supported by the National Science Foundation under award 0619060.

The National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant Number 0072909), the University of California, and the UC Santa Barbara campus

Collaborators:

NCEAS (UC Santa Barbara); UC Davis; UC San Diego (San Diego Supercomputer Center); CENS, University of California, Los Angeles; OPeNDAP; Oregon State University